Structural studies on proteins with potential therapeutic value

Thiyagarajan, Nethaji

Award date:
2006

Awarding institution:
University of Bath

STRUCTURAL STUDIES ON PROTEINS WITH POTENTIAL THERAPEUTIC VALUE

Submitted by Nethaji Thiyagarajan
for the degree of Ph.D.
of the University of Bath
2006

COPYRIGHT

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Botulinum neurotoxins are the most potent toxins known and are produced by the anaerobic bacteria of the genus *Clostridium.* These toxins are the causative agent of botulism (in human, animals, and birds), a neuroparalytic illness. Despite their potent toxicity botulinum neurotoxins are used in therapeutics and cosmetics. There exist seven antigenically distinct serotypes of botulinum neurotoxins (A-G). All botulinum neurotoxins have a similar domain arrangement - catalytic domain, transmembrane domain, and cell-binding domain. The toxin is non-toxic in the absence of cell-binding domain and such non-toxic fragments were highly researched for use in therapeutics. Structural characterisation of the non-toxic fragments would help in harnessing their effect as a effective pharmaceutics. Crystallisation studies and successes on the non-toxic fragments of botulinum neurotoxins and their first X-ray diffraction data are reported.

The entry of botulinum neurotoxin toxin into the host neurons is still not completely understood. Botulinum neurotoxins are thought to be internalised by receptor-mediated endocytosis after the toxins bind to the neuronal cell surface. Here we report the crystallisation trials and structural studies to obtain botulinum neurotoxin B in complex with its receptor synaptotagmin II.

Erythrina cristagalli lectin (ECL) is a carbohydrate binding protein specific for galactose and *N*-acetyl galactosamine and is a biological dimer. ECL in conjugate with botulinum neurotoxins non-toxic fragments was used to retarget cells as a therapeutic in the treatment of chronic pain. Understanding the structural mechanism of ECL will aid in using it as a drug delivery vehicle. It is thought that the dimer formation between two monomers is mediated by *N*-linked glycosylation ('hand-shake' model). But, our recent study and results tells that the *N*-glycosylation is independent of the dimer formation indicating that the functional properties of ECL are dictated at its primary sequence itself.

Ribonucelases are enzymes that catalyse the cleavage of phosphodiester bonds in RNA. Bovine pancreatic ribonuclease A (RNase A) and its homologues possess different properties and functions like angiogenesis (angiogenin), anti-viral and anti-bacterial (eosinophil-cationic protein (ECP) and eosinophil derived neurotoxin (EDN)) properties in living systems. An effective inhibitor of angiogenesis could inhibit the formation of new blood vessels in tumour cells, and inhibitors for ECP and EDN could be used to regulate the eosinophilic content of these proteins during inflammation processes. Here we report the biochemical and structural studies on binding of four different inhibitors to RNase A.

Acknowledgements

I would like to thank my supervisors Professor Ravi Acharya and Dr. John Chaddock for their help and guidance during my postgraduate research study at Department of Biology and Biochemistry, University of Bath and Health Protection Agency, Porton Down. I would also like to thank Drs. Philip Marks, Peter James, Lindsey Durose, Rob Fretwell, Jonathan Wayne, Mr. Roger Ling, Mr. Jeremy Whittles, Mr. Owen Wells, and Mrs. Bindu Mohan for their help and support during my work at HPA. I would like to thank University of Bath and Health Protection Agency for the post-graduate research. I would also like to thank Sir Richard Stapley Trust, UK for financial aid during my research. I am thankful to the station scientists at the Synchrotron Radiation Source, Daresbury Laboratory, UK for their support during X-ray data collection.

I would like to thank our collaborators Professor Ronald Raines, University of Wisconsin-Madison, USA for providing the ribonuclease inhibitors, and Dr. Thomas Binz, Dr. Andreas Rummel, Institut für Biochemie, Hannover, Germany for providing pure botulinum neurotoxins protein for structural studies.

I would also like to thank my colleagues past/present in the laboratory Drs. Shalini Iyer, Matthew Baker, Daniel Holloway, Stephen Prior, Hazel Corradi, Mr. Kenneth Holbourn, Ms. Eirini Mitsiki, Ms. Haryati Jamaluddin, Ms. Konstantina Kazakou and Mr. Balasubramanian Dhandayuthapani for their support and advice during my research.

I am grateful to Professor Ravi Acharya, Dr. Vasanta Subramanian and my friends in the lab Shalini, Eirini, Konstantina, and Balu for their moral support and making me feel at home.

I am thankful to Mr. Andrew Greenwood and his family for their support during my stay in Salisbury.

Finally, I would like to thank my parents, relatives and friends Geethalakshmi, Jayaprakash, Premalatha, Arunkumar, and Ethayathulla in India for their support.

Contents

Chapter 1

Protein Crystallography

Introduction

Living forms are classified as Viruses, Prokaryotes and Eukaryotes (Appendix I). The building blocks of all living forms are cells except viruses, which are made up of RNA (ribonucleic acid) protected by protein core. The functional units of cells are proteins (Appendix I), which are polymers of amino acids. Proteins play a very important role as transporters, catalysts, regulators, sensors, and in host defence. The function of a protein is defined by its structure as dictated by its sequence. Hence, structural characterisation of proteins is very important to understand the chemical nature of the protein and the mechanism of its function in a living cell. There are many methods available to determine the tertiary or three-dimensional structures of proteins, experimental methods include X-ray crystallography, Nuclear Magnetic Resonance spectroscopy (NMR), Electron microscopy, Neutron scattering, and γ-ray resonance scattering.

Bioinformatics or computational biology applies the knowledge gained from the experimental studies to predict the tertiary structure of the protein from its primary sequence. This method of predicting three-dimensional structures of proteins is called homology modelling. The principle is that sequence determines the structure of a protein. Hence, more similar sequence should have a similar structure, which is the basic criterion for homology modelling. Yet, homology models are only models and are biased on the template used and cannot define the nature of arrangement or conformation of the protein molecules as defined by the experimental methods.

1.1 X-ray Crystallography

X-ray crystallography is the most superior of all techniques known today to determine the three dimensional structures of macromolecules. Macromolecular crystallography has become an integral and important part of modern biology. It finds its use not only in structure determination of

molecules, but also guide efforts to chemically modify (Appendix I) the molecules to aid in understanding their functional/biological role in an organism.

After the discovery and report by W. C. Röntgen in 1895, that X-rays can affect a photographic plate and failed to possess wave nature i.e. X-rays are highly penetrating radiations and do not obey laws of reflection, refraction and/or diffraction, it was Max von Laue in 1912, first to beam the X-rays on to a Zinc sulphide crystal placed between the X-ray source and the photographic film. He observed a diffraction pattern on the photographic film similar to the interference phenomenon of light from an optical grating element. This pattern contained a regular array of spots, which were later called as *Laue Spots*. This phenomenon supported that X-radiations possess wave property and diffracts in the wavelength range of 0.1 Å – 100 Å.

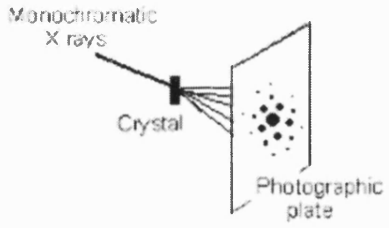


Fig 1.1 A schematic representation of Laue spots.

1.2 Bragg's Law (Bragg 1913a; Bragg 1913b)

W. L. Bragg noticed the similarity of diffraction to ordinary reflection and deduced a simple equation treating *diffraction* as *reflection* from planes in the lattice.

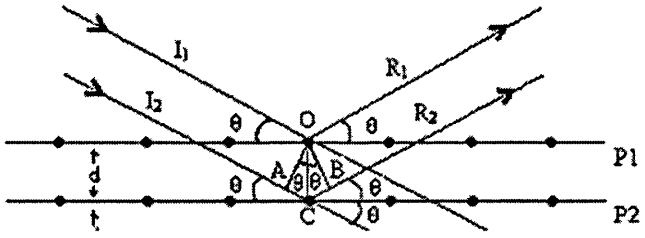


Fig 1.2 Bragg's law

Let I_1 and R_1, I_2 and R_2 be the incident and reflected (diffracted) X-rays (of wavelength λ) from the planes *P1* and *P2* respectively. Planes *P1* and *P2* are separated by a distance *d*. The incident and the reflected beam are inclined at an angle θ to the planes *P1* and *P2*. According to the law of diffraction, for a two waves to be in phase the path difference must be an integral multiple of the wavelength. Thus, *AC+CB* the path difference between I_1, R_1 and I_2, R_2 should be an integral multiple of wavelength for the two waves to travel in phase.

$$AC+CB = 2AC = n\lambda$$

$$AC = d \sin\theta$$

Thus, $n\lambda = 2d \sin\theta$ ---- Bragg's law.

The unit of wavelength λ is Å and $1Å=1x10^{-8}$ cm

1.3 Crystal System and Symmetry

A crystal can be defined as a highly ordered big block containing a uniform array of the atomic motif repeating in three-dimensional space.

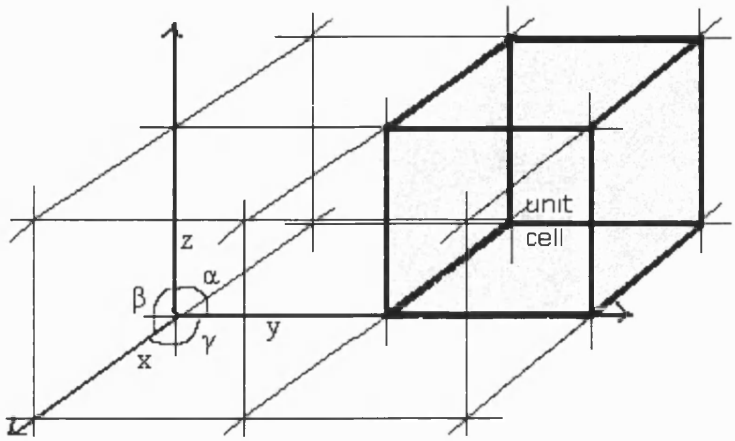


Fig 1.3 Diagramatic representation of unit cells in a crystal.

Crystals exist in different forms (Table 1.1). Based on their type and property they are divided and grouped under a common crystal class or system. The seven types of crystal systems given in the Table 1.1 are in the increasing order of symmetry based on the number of independent parameters. The lengths of the unit cell edges are designated as *a*, *b* and *c* and the angle between the axes *b* and *c* is *α*, *a* and *c* is *β*, *a* and *b* is *γ* (Fig 1.3). In addition to these seven *primitive lattices* Bravais discovered another seven *non-primitive lattices*. These seven primitive lattices and seven non-primitive lattices are together called *Bravais lattices*. Primitive lattices are those crystal systems that contain one point at each corner of the unit cell and are designated by letter P. Non-primitive lattices have points in the center of the unit cell faces or unit cell itself in addition to the points at the corner of the unit cell and are classified as face centered (A, B, C, and F), and body centered (I).

Table 1.1 Seven crystal systems of Primitive lattices.

Crystal system	Number of degrees of freedom	Parameters
Triclinic	6	$a \neq b \neq c$; $\alpha \neq \beta \neq \gamma$
Monoclinic	4	$a \neq b \neq c$; $\alpha = \gamma = 90^0$; $\beta > 90^0$
Orthorhombic	3	$a \neq b \neq c$; $\alpha = \beta = \gamma = 90^0$
Tetragonal	2	$a = b \neq c$; $\alpha = \beta = \gamma = 90^0$
Trigonal		
Rhombohedral lattice	2	$a = b = c$; $\alpha = \beta = \gamma \neq 90^0$
Hexagonal lattice	2	$a = b = c$; $\alpha = \beta = 90^0$; $\gamma = 120^0$
Hexagonal	2	$a = b = c$; $\alpha = \beta = 90^0$; $\gamma = 120^0$
Cubic	1	$a = b = c$; $\alpha = \beta = \gamma = 90^0$

1.3.1 Point Groups

This is a symmetry group operation that can be used to define the symmetry (arrangement of atom or molecule) of a crystal using the symmetry operators like rotation, mirror plane, inversion and center of symmetry. There exist 32 point groups.

1.3.2 Space Groups

Space group operation is a representation (arrangement of atom or molecule) of a three dimensional crystal, using all symmetry elements including glide planes and translations present in the crystal. 230 unique space groups (listed in International Tables of Crystallgraphy, Volume A (Hanhn et al. 1987)) are derived from 32 point groups and 14 Bravais lattices.

1.4 Protein Structure Determination by X-ray Crystallography

Structure remains the ultimate answer to define and describe the functions of proteins or enzymes and their chemical nature. The events in the determination of 3D-structure by X-ray crystallography are shown as a flowchart in Fig 1.4.

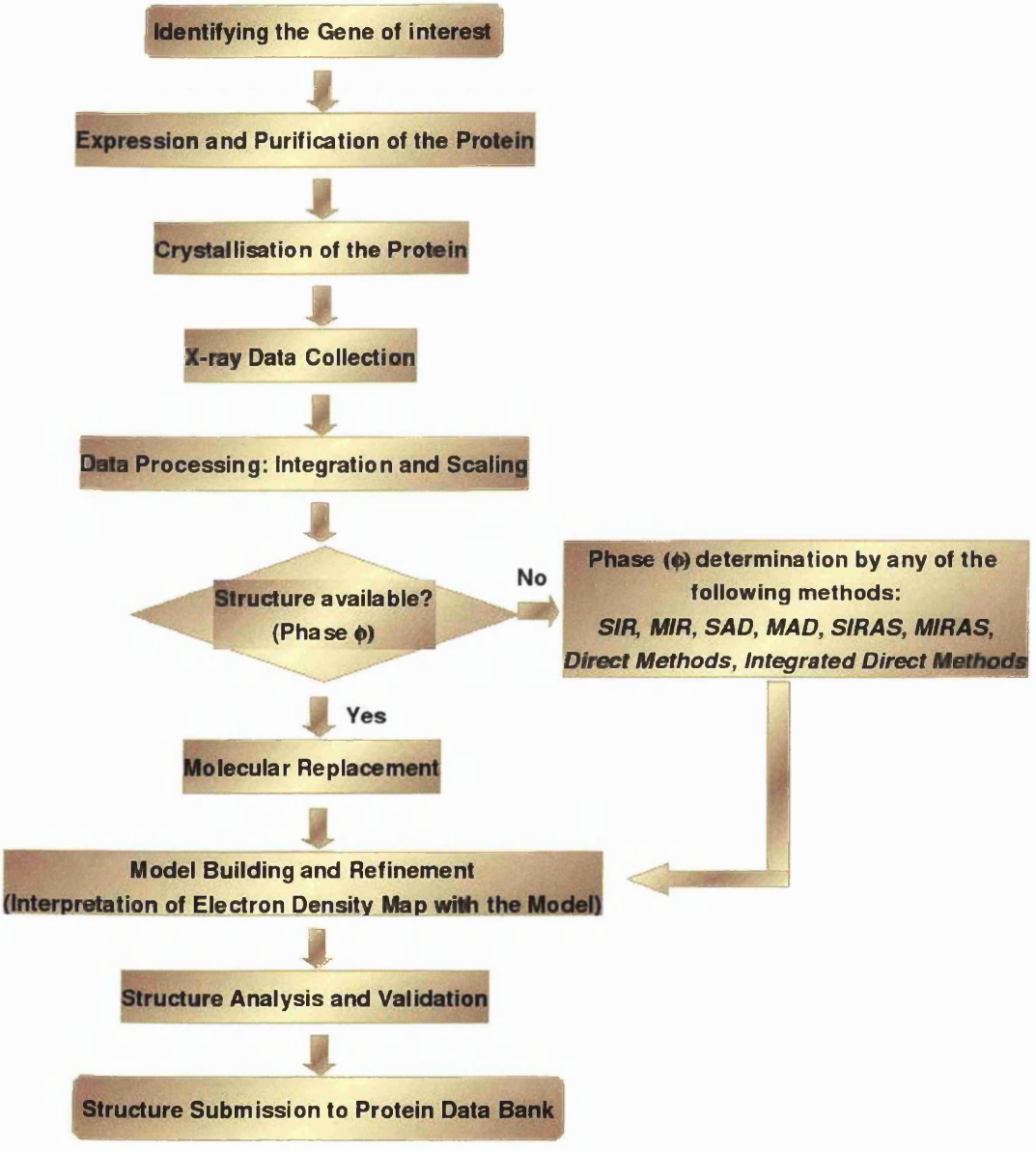


Fig 1.4 A flowchart representation of typical steps in molecular structure determination by X-rays.

The pre-requisite for X-ray crystallographic diffraction studies is to have the crystals of the protein of interest. Proteins can either be obtained from the native/natural source if possible or over expressed in different systems depending on their availability and ease of production. Pure protein is then screened for crystallisation trials, which will be further used for diffraction studies. In cases where obtaining pure protein from wild source is rather difficult or the protein is from mammalian cells, the target protein is expressed in a foreign host where one has many choice of selection.

1.4.1 Protein Expression

Protein expression *in vitro* is the choice of most researchers to obtain the protein of interest in large quantities. The important criteria for *in vitro* expression in selection of host are:

- The protein of interest is non-toxic or harmless to the foreign host.
- Protein is expressed in soluble form. Few cases where expression is in insoluble form (aggregates) selection of different fusion protein/tags or a different host should be considered.
- In case of poor expression one can consider different bacterial strains and also different expression vector system.
- Preferably as a non-interacting species to any other protein in the host.
- Capabilities of the host to over express the protein of interest.
- Should undergo post-translational modifications (as required for structural studies) as in the native source to become a fully mature product.

Several systems such as bacteria, insect cells, drosophila cell lines (Yamaguchi et al. 1965; Winslow et al. 1989; Wang et al. 1993; Torfs et al.

2000), yeast (van den Berg et al. 1990; Fleer et al. 1991; Rocha et al. 1996; Tokunaga et al. 1997; Walsh and Bergquist 1997; Walsh et al. 1998), mammalian cells (Laemmli 1970; Graham and van der Eb 1973; Wigler et al. 1977; Xu and Beckett 1994), and cell-free systems can be a choice of host selection to be used for the production of target protein using the recombinant method. *Escherichia coli* is the most commonly used *in vitro* expression system for most of the proteins. The advantage of use of *E.coli* in expressing the recombinant protein is that the whole genome information about *E.coli* is well studied and characterized and the bacterium has a high growth rate compared to any other systems. In cases where posttranslational modification of the target protein is not achieved in conventional bacterial expression system like *E.coli*, yeast or eukaryotic cell lines are the other choices of selection to express the protein of interest. The next step is to purify the over expressed protein using any of the available protein purification methods.

1.4.2 Protein Purification

Protein purification, whether from native source or by recombinant method is a challenging task. Protein purification can be single-step purification, but most of the time requires more than one step to obtain the final pure product. Protein purification normally involves chromatographic techniques and many different chromatographic methods are available. One needs to be selective in the chromatographic method used or in the combination of different chromatographic techniques for purification of any protein. There is always some loss of protein of interest in each purification step, so one has to be selective and limit the number of steps involved in purification but not at the cost of purity. This could be achieved by expressing the target protein along with affinity fusion tags/proteins such as Glutathione S-Transferase (GST) (Smith and Johnson 1988), *poly*His (Studier and Moffatt 1986; Rosenberg et al. 1987; Studier et al. 1990), Maltose Binding Protein (MBP) (di Guan et al. 1988; Maina et al. 1988), Calmodulin (Aslanidis and de Jong 1990), Biotinylated fusion tags (Samols et al. 1988; Cronan 1990; Wilson et al. 1992;

Xu and Beckett 1994), Intein-Mediated Purification with an Affinity Chitin-binding Tag (IMPACT) (Xu et al. 1993; Chong et al. 1996; Xu and Perler 1996; Chong et al. 1997; Chong et al. 1998a; Chong et al. 1998b), polyHis•Strep (Schmidt et al. 1996; Skerra and Schmidt 2000; Smyth et al. 2000)etc. Commonly used chromatographic techniques are Affinity chromatography, Ion-exchange chromatography, and Size-exclusion chromatography.

Affinity chromatography such as Ni-NTA exploits the principle that the target protein binds/chelates reversibly to the Ni-NTA charged column and eluted with different concentration of another ligand (usually imidazole) that possess higher affinity to the protein. Ion-exchange column chromatography separates proteins based on the charge carried by the protein. In this method all proteins including impurities that are bound to the ion-exchange column are separated by a salt gradient elution. Size-exclusion chromatography separates the proteins according to the molecular weight of the protein. Typically size-exclusion chromatographic technique is the final polishing method applied to get the purest sample. An SDS-PAGE is run after each chromatographic technique used or after each step used in the protein purification processes to ensure that the protein purified is pure. Biophysical (Circular Dichroism (CD) and Mass spectrometry) and biochemical characterisation should be carried out to confirm that the protein purified is fully mature and functionally active.

1.4.3 Protein Crystallisation

Crystallogenesis is a process of growing proteins as crystals. Almost all proteins are very selective about the chemicals and reagents used to aid them in crystallogenesis. There are two important phases in the process of crystal growth - *nucleation* and *growth.* Nucleation is a process of aggregation of similar molecules in the crystallisation medium, which results in

supersaturation of the protein molecules followed by growth phase. In the growth phase the crystal grows in three-dimensions. Depending on the polarity and arrangement of the protein molecules the crystals grown will be of different shape. Macromolecular crystals are not solid crystals but are intermediate between fluid and solid and are called crystalline solids. Protein crystals have solvent channels in between neighbouring molecules of the crystal.

Crystals can be grown by many different methods like Hanging drop method, Sitting drop method, Batch method, Sandwich method, Dialysis button method, Gel and Microbatch method. Many factors that influence the crystal growth are purity of the sample, pH, salt, temperature, stability, etc. The following diagram explains the phases in the crystal formation by vapour diffusion method.

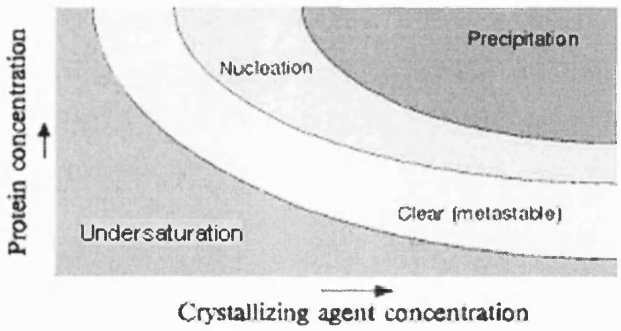


Fig 1.5 Phase diagram of crystallisation process

Hanging drop method: The most commonly used method to grow protein crystals is by the hanging drop technique (Fig 1.6). One of the main advantages of this method is that it requires less protein than any other methods. A few micro litre of protein solution is mixed with an equal amount (usually 1:1) of reservoir solution containing the mother liquor and placed as a drop on the cover slip. Crystallisation is achieved by supersaturation. A thin layer of grease seals the cover slip on the well containing the mother liquor. The most commonly used plates are 24 well Linbrow plates.

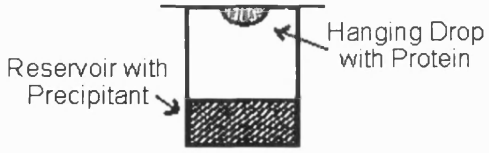


Fig 1.6 Hanging drop method.

1.5 X-ray Diffraction Data Collection

Any crystals grown should be exposed to X-rays to record the diffraction pattern, which provides the source of data for structure solution. The X-ray diffraction data collection depends on the choice of the crystal and the wavelength of the X-ray used.

1.5.1 Choice of Crystals for X-ray Diffraction Studies

The diffracting property of a crystal depends solely on the quality of the crystal grown. Not necessarily all crystals grown under the same condition and/or in the same drop will exhibit the same quality of diffraction. Hence, choice of crystal selection is important in diffraction data collection. One cannot come to a conclusion about the choice of a crystal by the look of it. Salts in the mother liquor might get crystallised instead of protein and it is very difficult to characterise them under the microscope. The only way to check the crystal quality is to shoot them with X-ray beam and record the diffraction strength. Some crystals might diffract strongly, some weakly, and others fail to exhibit diffraction. Salt crystals can be easily differentiated from protein crystals by looking at their diffraction pattern.

1.5.2 X-ray Sources and Detectors

X-rays are produced by bombarding high velocity electrons on to a metal target. X-radiation lies in the region between the ultraviolet rays and gamma rays in the electromagnetic spectrum and is of the order of wavelengths 0.1 Å – 100 Å. A commonly used metal in the production of X-rays is Copper for its advantage in efficient recording of the diffraction data and suffers less absorption while passing through the medium or air. Other metals used are Molybdenum, Chromium, Iron, etc. For a long time in-house X-ray sources have been used to generate X-rays (eg. a rotating anode). However, the intensity of the emitted beam is very low as it is directly proportional to the number of electrons emitted at any given time and the images have to be recorded for a very long time. Synchrotron Radiation Sources (SRS) are routinely used for X-ray diffraction data collection and the X-rays produced are of more intense than the in-house source and the diffraction measurement requires less time.

Use of synchrotrons in X-ray diffraction studies has several advantages of which reduction in the duration of data collection from days to hours is one of the important factors. X-radiation generated at synchrotrons is of very high intensity and greater magnitude of about 2 GeV. The SRS offers a wide range of wavelengths for different experiments extending from the infrared to hard X-rays to different techniques including X-ray diffraction, X-ray spectroscopy, small-angle/wide-angle scattering, soft X-ray spectroscopy, photoemission, imaging, nanoscience, time-resolved X-ray spectroscopy, etc. In synchrotrons the electrons are amplified in a booster and are fed into the high-energy storage ring where the electrons travel at relativistic speeds of light and gain high momentum. The X-ray diffraction experiments are carried out in a hutch positioned tangential to the storage ring where the circulating electron beam is released from the high-magnetic field. Charge Coupled Device (CCD) detectors are superior to image plates as they have a very good spatial resolution, good noise reduction, and shorter read-out time. CCD detectors are now commonly used in in-house source and in SRS.

1.5.3 X-ray Data Collection

The next step in structure solution is *X-ray diffraction data collection.* Diffractrometers are instruments used to collect X-ray diffraction data. The X-rays are beamed on to the crystal mounted on the goniometer. The goniometer is a part of the diffractometer where the crystal is mounted and positioned in space to be in the center of the X-ray beam. There are many parameters such as distance between the crystal and detector, crystal oscillation range ($\Delta\phi$) of the crystal about rotation ϕ-axis, the number of images to collect to get a complete dataset, and exposure time that one should consider for data collection. The diffraction range depends on the crystal to detector distance following Bragg's equation

$$d = \lambda/(2 \sin \theta)$$

The major difficulties in intensity measurements that arise are: (a) The scattering of radiation other than Bragg reflection, (b) The intensity of the beam, (c) The boundaries of the reflection must be correctly defined irrespective of the spot size, shape and intensity, (d) The diffraction spots are always accompanied by other scattered radiation, called as background, which must be adjusted while data processing. This might be due to diffuse scattering of X-rays of all wavelengths by all objects in the beam, including air and crystal.

The molecular information that one can obtain from X-ray diffraction data is solely dependent on the resolution (wavelength λÅ) of X-radiation used.

The concept of Resolution

Resolution is defined as the distinct interpretation of two closely spaced spots (atoms by X-rays) on the diffraction image. Thus, the details of the structure is proportional to the resolution used (Blundell and Johnson 1976; Acharya and Rees 1995). In terms of X-rays resolution is defined as the wavelength λ of the incident beam. The information in Table1.2 gives an idea about what one might expect at the given resolution.

Table 1.2 Details of structure one might expect at given resolution.

Resolution (Å)	Details of the structure
6.0	Outline of the molecule, secondary structural features (helices, beta strands) can be identified.
3.0	Course of the polypeptide chain can be traced and topology of the folding can be established. With the aid of amino acid sequence, it is possible to place the side chains within the electron density map.
2.0	Main chain conformations can be established with great accuracy. Details of side chains, some bound water molecules, ligand, and metal ions can be identified.
1.5	Individual atoms are almost resolved. It is possible to assign large part of the solvent structure.
1.0	Hydrogen atoms may become visible. Multiple conformations for some of the side-chains can also be seen.

The concept of Reciprocal Lattice and Ewald Construction

From Bragg's law

$$\sin \theta = (n\lambda/2)*(1/d)$$

sin θ is directly proportional to *1/d*, the interplanar spacing in the crystal lattice. When crystals diffract X-rays, closer diffraction pattern is observed for lattices with larger *d* in real space and wider diffraction pattern for lattices with smaller *d*. Thus, the diffraction patterns observed are in reciprocal space

(1/d). Not all the points on the lattices contribute to any observed diffraction at a time. Only those lattice points that lie within the *limiting sphere* contribute to the diffraction. Limiting sphere is a sphere in three-dimensions constructed with radius $1/\lambda$. All lattice points that fall on the circumference/on the surface of the limiting sphere diffract X-rays. This construction is called *Ewald sphere* (Fig 1.7). The crystal is rotated about its rotation φ-axis and the diffraction data are collected. Irrespective of the degree of rotation the lattice spacing remains the same throughout during the process of X-ray diffraction data collection.

Consider a circle (Fig 1.7) with centre C and radius AC = 1/λ, where A is any point on the circumference of the circle where the incident beam I meets. O is the direct emerging beam and is collinear with the incident beam I. As per *Ewald's construction* all points that lie on the circumference of the sphere of reflection contribute to diffraction phenomenon. D, E, F, G, and H are the points from different and/or same lattice planes of the crystal C that lie on the circumference of the limiting sphere. R_1, R_2, R_3, R_4, and R_5 are the reflected beams when Bragg's law is obeyed.

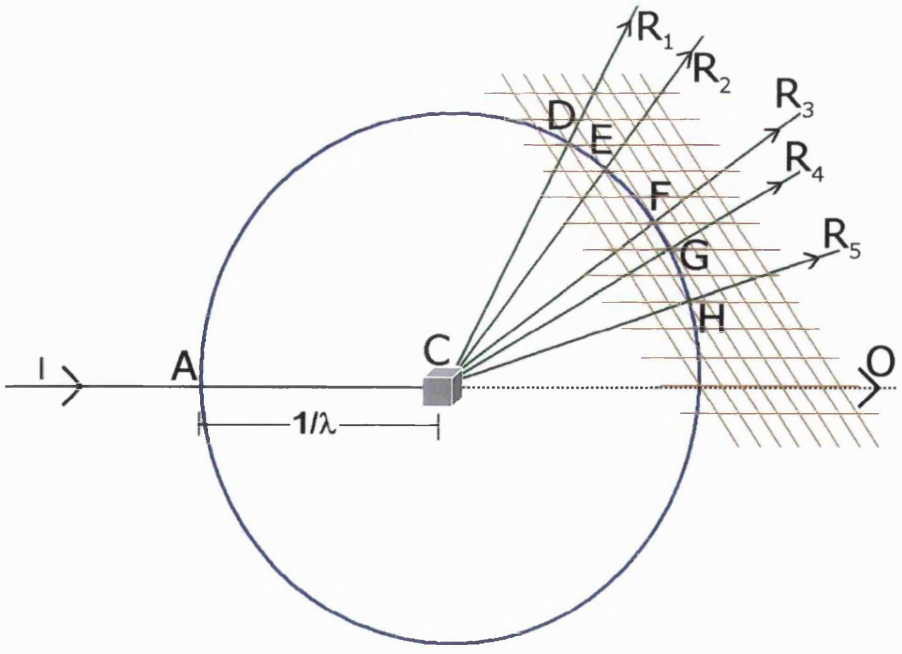


Fig 1.7 Ewald construction

Radiation damage

Radiation damage to the crystal in X-ray beam is an unavoidable situation during the process of X-ray diffraction data collection. This arises due to several factors such as evaporation of solvent molecules in the crystal, and heat produced by the incident X-ray beam (quantum effect). Radiation damage leads to disorderness of the crystal lattice resulting in low-resolution diffraction data and/or loss of resolution and diffraction. This problem can be overcome to certain extent by collecting the data at cryo temperatures (usually at 100K), where a stream of nitrogen gas is focussed on the mounted crystal. Cryo temperatures could freeze solvent molecules in the crystal leading to ice formation on the crystal resulting in poor quality diffraction data. The process of ice formation at cryo temperatures could be successfully avoided by soaking the crystal in cryo protectants (glycerol, MPD, etc.) prior to data collection.

1.6 Data Processing, Data Reduction and Scaling

The raw data from X-ray diffraction data collection or Intensity data collection contains all information about the crystal, symmetry and the structure of molecules in terms of observed intensities only. One need to know the position of atoms/molecules called as *phase* in order to define the structure, which is discussed later in the section '*Phase problem*'. Firstly, one has to extract all information from the raw data by carefully examining the diffraction images. Secondly, the obtained information is to be scaled to contain unique set of reflections that will define the molecule.

1.6.1 Data Indexing and Processing, and Integration of the Raw Images

All the observed intensities (spots) on a diffraction image must be defined a position with reference to other observed intensities. This process of defining

all the spots a defined position is called *indexing*, which is the first step in X-ray diffraction data processing. All cell dimensions like the axes lengths, angles are calculated and the possible lattice type (space group) is selected using the conditions in the Table 1.1. For a long time scientists were doing the indexing with their eyes with the help of simple devices to measure the intensities, cell dimensions, etc. Development in detectors, computational crystallography and automation has given rise to new software suites, which are capable of indexing the images automatically (Otwinowski and Minor 1997a). The indexed images are to be refined for the defined positions for the unique space group. This step is very important as it gives information about the choice of selection of space group. There are many parameters one has to take into account while the process of indexing and processing like intensity spot radius, the background of the spots – very important to define the boundary of each spots in order to avoid the overlap of neighbouring spots, the shape of the spots, and the spread of the spots in a diffraction image (mosaicity). Mosaicity is an important parameter, which defines the quality of the crystal and data collected. Greater mosaicity means the high degree of disorderness in the crystal lattice. One has to refine all the parameters carefully to get the most out of the diffraction images and to use all the data collected.

1.6.2 Data Reduction and Scaling

Data reduction and scaling is the final step in X-ray diffraction data processing. The *'n'* number of images collected for a crystal by rotating about its rotation φ-axis contains partially observed reflection of the previous image and the following image. Also some of the observed reflections will be the same in a couple of images. This is because when the crystal is rotated the same lattice points will fall on the limiting sphere after a few degree of rotation and Bragg's law is obeyed. This redundancy in the data should be reduced and scaled to a non-redundant dataset where all the sets of reflections after scaling are brought to a common scale and are unique and contain all the

information about the crystal. Other statistical information obtained from scaling is an overall estimate of mean thermal motions/vibrations for the molecule in the crystal space.

1.7 Phase Problem

What we actually record during X-ray intensity data collection?

Is there any information lost during X-ray data collection and if, how can it be obtained?

When we know the symmetry, unit cell dimensions and unit cell volume what are the missing parameters that would define the position of molecules in the unit cell?

These questions look simpler but involve lots of vector mathematics and statistics. The position of a molecule in space is defined by its direction and magnitude to a reference point. The crystallographically available data consist only of the structure factor magnitudes ($|F_{hkl}|$) in the form of intensities/amplitudes and not their positions. This is called the *phase problem*. Thus one has to obtain the phase information lost during X-ray intensity data collection in an indirect way before reconstructing the three-dimensional image from the diffraction pattern.

By definition, structure factor is a vector quantity, which is the resultant of *N* waves scattered by the *N* atoms in the unit cell. Each of these waves has amplitude proportional to f_j, the atomic scattering factor and contains the information about phases. Structure factor in simple exponential form is given by the formula,

$$F_{hkl} = \Sigma f_j e^{2\pi i h.x}{}_j$$

Where, F_{hkl} => Structure factor of the atom

f_j => atomic scattering factor of the jth atom

hkl => Miller indices

1.8 Methods to Determine Phases (φ)

There are many methods by which one can derive the phases (φ) of the molecule.

1.8.1 Molecular Replacement Method

Molecular replacement method (Rosmann and Blow 1962; Rosmann 1990) is the first choice of all the structural biologists when there is a homologous structure available. Homology between the observed data and available structure is determined at the primary structure level. Successful calculation of approximate phases for an unknown structure from a known molecule, which shares homology greater than 20 % to the target sequence, has been achieved using this method. The principle is to find the relative orientation (rotation) and position (translation) of the unknown in space by performing rotation and translation operations on the known phases irrespective of the space group of the known molecule. MOLREP (Vagin and Teplyakov 1997), AMORE (Navaza 1994), and Phaser (Storoni et al. 2004) are a few of the most commonly used programs of this technique.

Patterson synthesis

Patterson function is applied or is integrated as a part of the molecular replacement method to find the orientation and position of the molecules in a unit cell. Patterson function is a vector representation of the observed intensities given by the equation

$$P(uvw) = (1/V)\sum |F(hkl)|^2 \cos[2\pi(hu+kv+lw)]$$

where, *uvw* are Patterson coordinates

Note: The equation has got only the *cosine* term and not the *sine* term which tells that the $|F|^2$ are phaseless quantities or phase equals zero degree.

Pattern synthesis thus produces a map of intra- and inter- atomic vectors. The intra-atomic vectors are self-vectors arising from the same molecule (vector distances will be smaller) and the inter-atomic vectors are cross-vectors arising between different molecules (vector distance will be larger) in a unit cell. For a molecule containing *N* atoms there will *N(N-1)* vectors/peaks in a Patterson map. Thus the relative orientation of the molecules is defined by the self-vectors while the position or translation is defined by the cross-vectors. Any Patterson function is centrosymmetric and the screw or translational symmetry are substituted with rotation and mirror plane symmetry. This is because for any pair of atoms in the molecule there are calculated vectors of opposite direction and same magnitude.

Steps in Molecular Replacement Method: (Fig 1.8)

1. Identification of the phase target (a target is the known structure) by sequence homology: Here one searches for any available structure in the Protein Data Bank with similarity (>20 %) in primary sequence.

2. Calculation of Patterson vectors using Patterson co-ordinates for the observed data (unknown) and the target (known).

3. Superposition of known and unknown: Involves two searches called rotation and translation.

$$\text{unknown } (\varphi) = R\ (\text{known } (\varphi)) + T$$

where, R => Rotation function, T => Translation function

a. To find the orientation of the Patterson vectors of the known with respect to the unknown in the cell. This is called rotational search. In total there are three rotational axes.

b. Once the optimal orientation is found, the position in the cell must be found. This is called translational search. In total there are three translational vectors one for each rotational axes.

4. The best fit is found by calculating the agreement brought by superposition method and the phases are assig unknown.

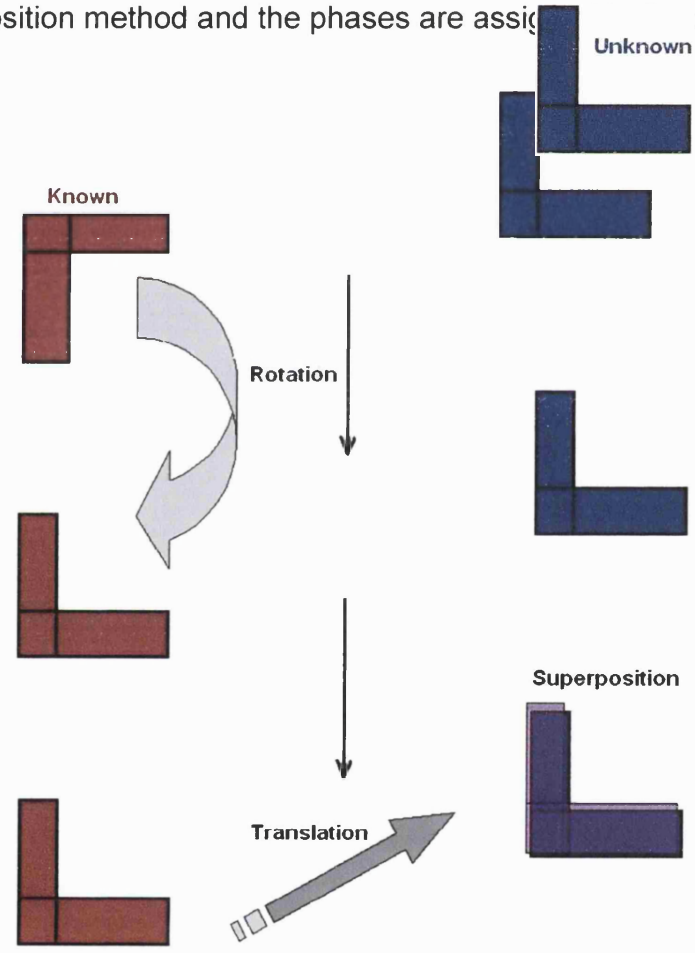


Fig 1.8 A diagramatic representation of Molecular Replacement Method. Known molecule – *brick red*; unknown – *blue*.

1.8.2 Isomorphous Replacement Method

Isomorphous replacement method (Green et al. 1954) exploits the principle that each atom in the molecule contributes independently to the scattering factor (amplitudes). A heavy atom (metal) bound to the molecule will contribute more to the scattering power than any other atoms in the molecule and is more prominently observed in the difference Patterson map. Two dataset, one for the native and the other for the derivative set (heavy atom soak) are collected and are called isomorphous when there is no observable difference in their lattice organisation after the introduction of a heavy atom. Concentration/selection of heavy atom should aim not to alter/drastically change the lattice properties of the crystal. If the introduction of heavy atom does not alter the rest of the structure called non-isomorphism, by constructing Harker sections, one can determine the phase (ϕ) information of the molecule. Usually more than one isomorphous derivative can be a better choice in determining the phase of the molecule, thus reducing the phase ambiguity. Single Isomorphous Replacement (SIR) method (Blow and Rossman 1961) and Multiple Isomorphous Replacement (MIR) method are used to determine the phases by this method.

1.8.3 Anomalous Scattering Method

Anomalous scattering (resonance scattering) of atoms is observed when the wavelength of X-radiation used is close to the absorbance edge of the scattering atom present in the crystal. Application of anomalous scattering method is amenable at synchrotron sources where one can tune the wavelength close to the absorbance edge of the anomalous scatterers present. Phasing with the help of anomalous scatterers in the X-ray diffraction data (Hendrickson et al. 1985) is similar to isomorphous replacement method, but taking the advantage that the conditions at which the scattering factor differs with the dispersion. Under normal conditions when non-anomalous scatterers are present, the reflections obey Friedel's law, i.e. $F_{hk\ l} = F_{-h-k-l}$. In

the presence of anomalous scatterers, the Friedel's law is violated, thus $F_{hkl} \neq F_{-h-k-l}$. Similar to SIR/MIR, anomalous scattering data collection can also be done using Single-wavelength Anomalous Dispersion (SAD) and Multiple-wavelength Anomalous Dispersion (MAD) techniques to determine the phases. However, the method has a drawback that the anomalous difference signal between dataset is sometimes too weak. The method can become more powerful when combined with isomorphous replacement technique and is called as Single wavelength Isomorphous Replacement and Anomalous Scattering (SIRAS) and Multiple wavelength Isomorphous replacement and Anomalous Scattering (MIRAS).

1.8.4 Direct Methods (DM) And Integrated Direct Methods

Direct Methods

Deriving phase relationships directly from the observed intensities to solve the phase problem is called *Direct Methods*. Phases are derived using symbolic addition methods. A few assumptions of direct methods are:

i) The structure factors are considered as normalized structure factors by the equation

$$|E_H|^2 = \frac{|F_H|^2}{\sum_{j=1}^{N} f_j^2}$$

where, E_H => Normalised structure factor

F_H => Structure factor

f_j => Atomic scattering factor

ii) The electron density is always positive.

iii) Each atom is considered as point/unitary atoms.

Direct Methods are successfully applied to small molecules data to find phases. The usage of Direct Methods to solve protein structures is reported by Karle in 1989. One needs very high resolution data with high redundancy for a successful use of Direct Methods to obtain phases. Attempts to solve protein structure by *ab initio* method of using Direct Methods were successful for proteins with less than 3000 non-hydrogen atoms. However, Direct Methods when integrated with any of the well-known techniques of phase determination (Isomorphous replacement (Giacovazzo et al. 2001), Anomalous dispersion, Patterson search) can give better estimates of phases.

1.9 Phase Refinement and Model Building

The final agreement between the observed diffraction data and calculated phases is brought by a process called *refinement.*

What do we actually do in the process of refinement and what is refined against what?

These are a few questions addressed in this section.

1.9.1 Model Building using Fourier Maps

The structure factor F_{hkl} is considered to be the sum of scattering of all the free electrons in a unit cell and contains the phase information of the molecule. Electron density is defined as the number of electrons per unit volume and is the *Fourier transformation* of structure factors given by the equation,

$$\rho(x,y,z) = (1/V)\ \Sigma_{hkl} F_{hkl}\ e^{-2\pi i h.x_j}$$

where, $\rho(x,y,z)$ => electron density over the volume element V

Once there available initial sets of phases for a fraction of the molecule the next step is to extend these phases (model building) with the knowledge of chemistry of the molecule. This is done with the help of electron density maps (Fourier maps). Different Fourier maps are calculated and used as a guide in model building. The rule of a thumb in calculating electron density maps is to use all the reflections measured, as the observed data is always true and will show up clearly in the electron density maps. Usually most of the model building is done with two kind of Fourier maps calculated:

Difference Fourier Maps: denoted as $\Delta F = F_O - F_C$, where the observed data and calculated phases are given equal weights. The importance and beauty of difference Fourier map is that it will clearly show any wrong substitution made during the process of model building.

Double Difference Fourier Maps: the observed data is given double weight ($2F_O$) in calculating double difference Fourier maps ($2F_O - F_C$) to guide through the process of model building. Because when $F_O = F_C$, no electron density is observed. Thus, the observed data is given more weight and used along with ΔF during the process of model building.

Interpretation of electron density maps is a time consuming and laborious task in model building. Developments in computation and visualisation tools have provided crystallographers with graphics software suites/tools that enable them to do model building on a personal computer. Model building accounts for all observed data by knowing what is in the crystal or what was crystallised, like protein atoms, solvent atoms, metal ions, and ligands/substrates if any present. At each stage of model building electron density maps are calculated and any difference observed is fitted into the density and taken into account for next round of refinement.

1.9.2 Refinement of the Final Model

What do we refine and try to achieve during the process of refinement?

Refinement is always followed by model building until,

- all the observed electron density is accounted for, i.e. addition of ligands and/or solvent atoms after accounting for all protein atoms. Addition of solvent atoms depends on the resolution of the data collected (Table 1.2).

- one finds closer agreement between the calculated (F_C) and observed (F_O) structure factors.

The calculated phases (F_C) are refined against the phases observed (F_O) during the process of refinement. There are several parameters that one has to take into account during the process of refinement of the model like the positional parameters (bond length, bond angle, torsion angle, and planarity) of the molecule, and temperature factor for each atom. This is done in two different ways strictly abiding the rules of chemistry, thus refinement can be either *constrained* or *restrained.*

In a constrained refinement the bond lengths (Appendix I), bond angles (Appendix I) are considered as rigid bodies whilst refining only the dihedral angles (Appendix I). By doing constrained one can reduce the number of parameters required for further refinement. In a *restrained* refinement the positional parameters are refined with reference to the energy. The variables are bond lengths, bond angles, torsion angles, thermal vibration factor (B-factor). Refinement comes to an end where there is no observable change in the *reliability index* called the *R*-factor after accounting for all the difference electron density observed.

R-factor is given by the formula,

$$R = \frac{\Sigma_{hkl} ||F_o| - k|F_c||}{\Sigma_{hkl} |F_o|} \times 100\%$$

where, k is the scale factor

h,k, and l are the reciprocal lattice points

R_{free} value:

Overfitting the model by accounting for all the electron density observed could relatively bring down the R-factor. In 1992 Axel T. Brünger (Brunger 1992; 1993) introduced a factor called *free R* factor in addition to R-factor, which is calculated in a similar way as the R-factor. R_{free} includes a very small amount of reflections called *test set*, which is not included in the process of refinement along with calculated structure factors. This *free R* relates to the absolute phase error in the model and any overfitting of data or even unjustified addition of solvent molecules to the model data is accounted by increases *free R*-values.

1.9.3 Structure Validation and Submission

Structure validation is a process where fine-tuning of the refined structure is done. There are several tools available to validate the solved structure. Ramachandran plot (Fig 1.9) (a φ-ψ plot for main chain atoms in a peptide plane) (Ramachandran et al. 1963) provided the first tool in validation to check the allowed and disallowed conformations present in the observed structure. Statistics states that more than 85% of the total residues should be in the allowed region of the Ramachandran plot. PROCHECK (Laskowski et al. 1993) is used to assess the quality of the structure after the final round of refinement, which gives details about the bond length violation, bond angle

violation, planarity of the molecule and also Ramachandran plot. WHATIF (Vriend 1990) is a program for structure analysis, MOLPROBITY (Davis et al. 2004) is a program to validate the structure and contacts in the structure. The final step after structure validation is submission to the Research Collaboratory for Structural Bioinformatics Protein Data Bank (www.rcsb.org).

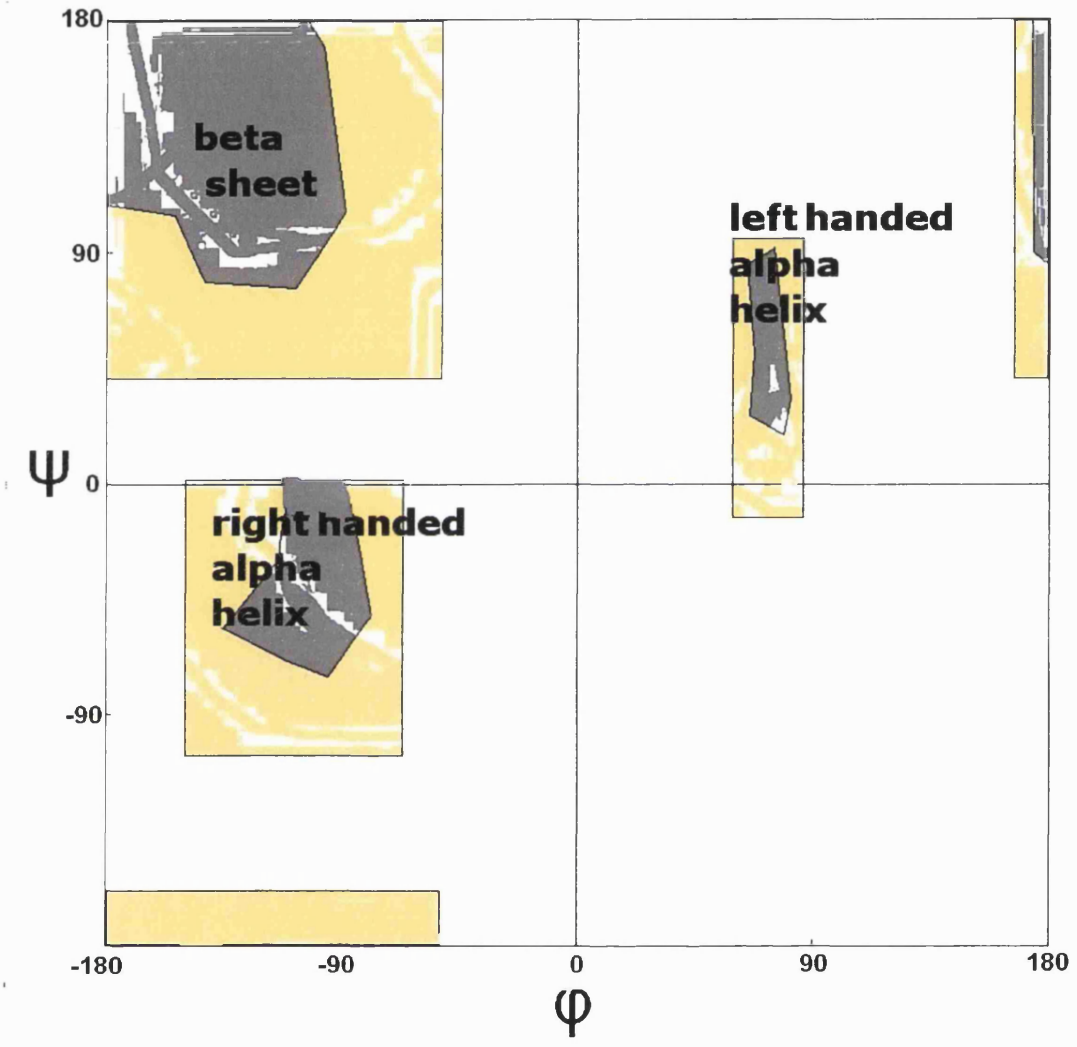


Fig 1.9 A diagramatic representation of Ramachandran plot. The unit of Φ and Ψ are in degree. Regions shaded in *grey* are the allowed regions of Ramachandran plot for respective secondary structure, while the regions shaded with *yellow* are the generously allowed regions of Ramachandran plot.

Recent Advances

Improvements in science and technology lead scientists to search for answers in many different ways. In the late 20th century several genomic projects deposited the complete genome of organisms including human genome into the public database. This opened the door to the researchers in different field of sciences and gave rise to new branch of sciences called Bioinformatics/Computational Biology, Structural Genomics/Proteomics, and Microarray technology (Appendix I). Also, X-ray crystallography has seen a significant development in the last decade that becomes very clear by looking at the growth rate of three-dimensional structures submitted in the Protein Data Bank.

Synchrotrons play a very important role and are an indispensable tool to structural biologists and to scientists in many other fields. The high intensity and high energy X-ray beam generated at Synchrotrons has reduced the X-ray diffraction data collection from days to hours and/or minutes. Time taken to solve protein crystal structures has decreased from months to days with the help of synchrotron radiation source, automation and robotics in this field. To date 3rd generation synchrotrons produce X-radiations of very high strength (3GeV - 7GeV) for scientific experiments. Synchrotrons not only generate X-radiations for structural biology, but also, ultraviolet radiations, LASER, X-radiations for medical scanning, etc.

Recent advancement and application of robotics in structural biology from crystallisation to data collection has made structural study a high-throughput one. Crystallisation robots can setup thousands of crystallisation trials in a day only using nanolitres of protein sample and can analyse the drops and report any success in them. In X-ray diffraction data collection robots are now used to mount the crystal, orient the crystal in the beam and to collect data. Automation in structure solution has been successfully tested on well known protein lysozyme by the application of robotics. Advances in Computational

Chemistry and Combinatorial Chemistry have opened the door to the drug companies towards rational design of drugs.

Though there observed significant development in all the fields the ratio of available three-dimensional structures to the gene sequences when compared is very low. This is because various techniques used to determine three-dimensional structure has bottlenecks to overcome. Availability of pure protein, crystallisation, and quality of the crystal and diffraction data are major checkpoints in X-ray crystallography.

In parallel with advancement in X-ray crystallography, developments were also reported in Molecular Biology and Protein Biochemistry. Novel techniques and developments in available methods for protein expression and purification are a boost to modern biology. Successes using different expression systems like yeast, viruses, insects, and mammalian cell lines are of very recent development in expression of mammalian proteins and many others where the traditional bacterial system fails. Automation and developments in protein purification techniques (section 1.4.2) and parallelisation of successive purification steps have limited the amount of manual intervention during the purification processes.

Chapter 2

Botulinum Neurotoxins

Introduction

Clostridial neurotoxins (CNT) are the most potent toxins known and are produced by the anaerobic bacteria of the genus *Clostridium* (Schiavo et al. 1994). This rod shaped anaerobic bacteria is commonly found in soil (Fig 2.1). The bacterial spores of *Clostridium botulinum* undergo hibernation until exposed to conditions that support their growth. *Clostridium botulinum*, *Clostridium butyricum*, *Clostridium barati* and *Clostridium tetani* are the causative agents of botulism and tetanus (Burgen et al. 1949; Kao et al. 1976; van Ermengem 1979; Simpson 1981). Botulism is a neuroparalytic disease caused by the neurotoxins produced by the anaerobic bacteria *Clostridium botulinum*. The symptoms of botulism include double vision, blurred vision, drooping eyelids, slurred speech, difficulty swallowing, dry mouth, and muscle weakness. If untreated, these symptoms may progress to cause paralysis of the arms, legs, trunk and respiratory muscles. The botulinum neurotoxin (BoNT) blocks the release of neurotransmitter at the nerve terminal (Humeau 2000; Lalli et al. 2003). There are three main kinds of botulism:

1. Food-borne caused by ingestions of the BoNT's.
2. Wound botulism – entry of the BoNT's through an open wound.
3. Infant botulism – caused by the bacillus spores in the infant intestine.

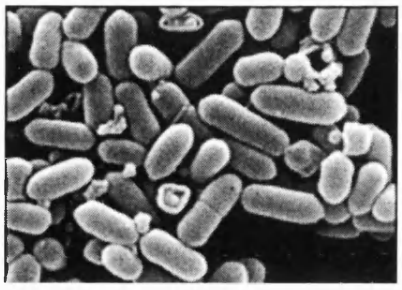


Fig 2.1 *Clostridium botulinum* – rod shaped anaerobic bacteria. (Image adapted from http://www.image.google.co.uk)

2.1 BoNT Structure

There exist seven different serotypes of BoNT's (A-G) (Lowenthal and Lamanna 1953; Kitamura et al. 1968; Kozaki et al. 1974; Ohishi and Sakaguchi 1975; Sugii and Sakaguchi 1975; Miyazaki et al. 1977; Iwasaki and Sakaguchi 1978). These toxins are synthesized as a single polypeptide chain (~150 kDa) (Beers and Reich 1968; Dasgupta and Sugiyama 1972a; b; Miyazaki et al. 1977). This single chain is subsequently cleaved to produce two chains: a single Light chain/catalytic domain (LC) (~50 kDa) and a Heavy chain (HC) (~100 kDa). The heavy chain is further divided into two subdomains – the N-terminal half (~50 kDa) called the translocation domain (H_N) and the C-terminal half (~50 kDa) called the ganglioside binding domain (H_C) (Fig 2.2a-b). A single disulphide bridge between residues Cys441-Cys450 links the HC and LC domains. The position of the cysteines in the amino acid sequence differs in different serotypes.

The H_C domain is further sub-divided into two subdomains (Fig 2.2a-b), which are composed of mainly β-sheet topology, joined by a α-helix. The jellyroll topology of the N-terminal sub domain of H_C domain is similar to the one found in lectins (Turton et al. 2002) and is well conserved among the CNT family. The H_N domain consists of a pair of long (105Å) amphipathic helices, which resembles the '*coil-coil*' motif as observed in some viral proteins (Turton et al. 2002). It is believed that the H_N domain might be involved in the translocation of LC domain into the cytosol by forming pores through the cell membrane. This domain contains a long flexible belt, which wraps the LC domain and occludes the active site of BoNTs, thus preventing access to the bound important zinc ion.

Several botulinum neurotoxin structures have been reported in the protein data bank. These structures includes light chain of serotypes A (PDB: 2G7K), B (PDB: 1F82), D (PDB: 2FPQ), E (PDB: 1T3A), F (PDB: 2A8A), and G (PDB: 1ZB7), binding domain of serotype B (PDB: 1ZOH), and whole toxin structure for serotypes A (PDB: 3BTA) and B (PDB: 1EPW) only.

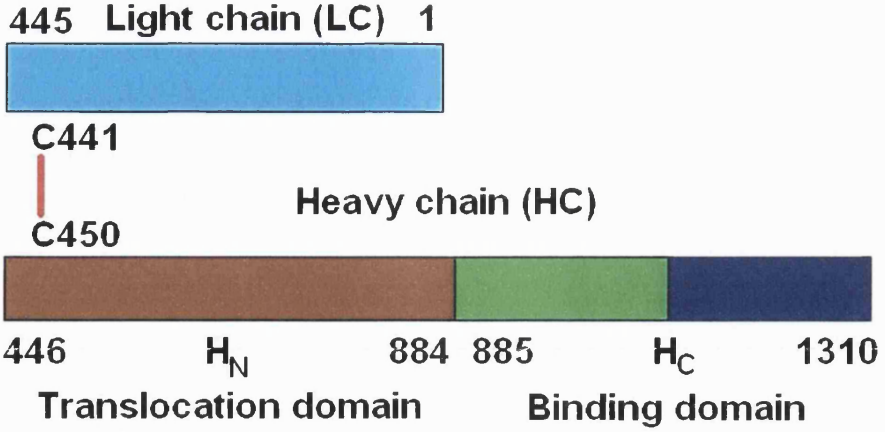


Fig 2.2a Schematic representation of domain organisation in *Clostridium botulinum* toxins.

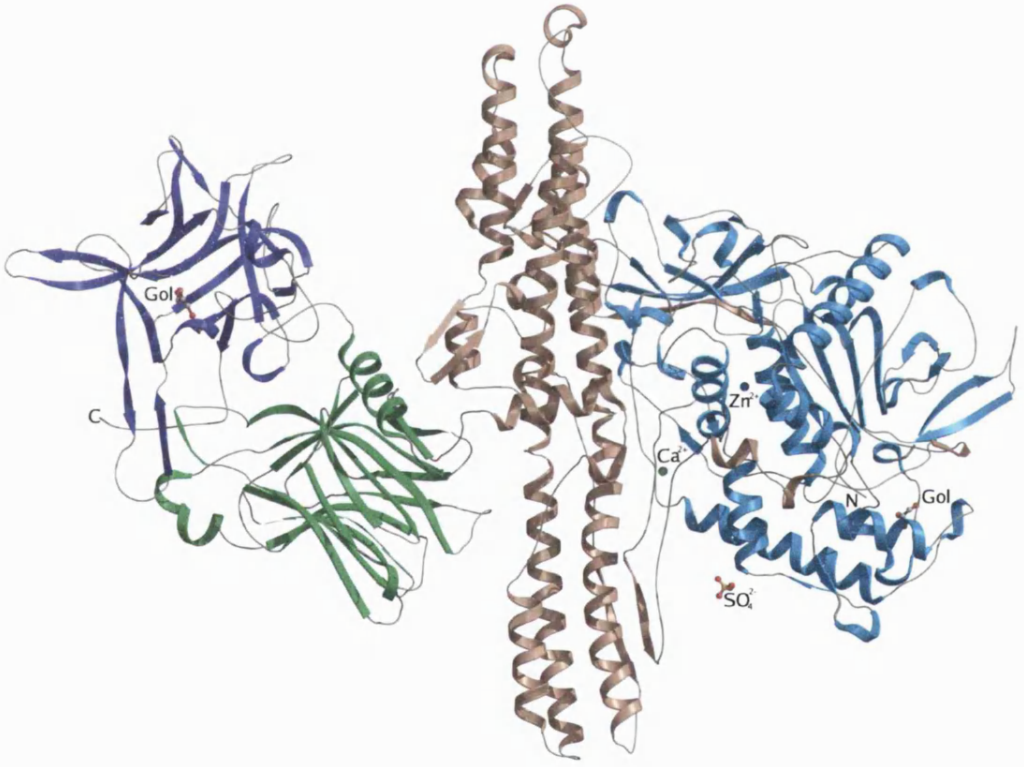


Fig 2.2b Ribbon representation of botulinum neurotoxin B (Thiyagarajan et al., unpublished results). Catalytic domain (light chain) – cyan; translocation domain H_N – brown; Binding domain – H_C: H_{CN} – green; H_{CC} – violet; Catalytic Zinc ion – blue; Calcium ion – green; sulphate ion and glycerol molecules are shown as ball-and-stick models.

2.2 Sequence Similarity of BoNTs and TeNT

CNT family consists of botulinum neurotoxin (BoNT) and tetanus neurotoxin (TeNT). CNTs share an overall sequence similarity of 30 % - 40 % within the family (Binz et al. 1990). Light chains of CNTs share a high sequence similarity within themselves and thermolysin, a zinc dependent metalloprotease (Holden et al. 1987). All CNTs possess three functional domains each involved in definite function (Fig 2.2b). CNTs could have evolved from a single ancestral gene (Ahnert-Hilger and Bigalke 1995) with high sequence similarity and conserved domain architecture (Lacy and Stevens 1999).

The LC domain contains a conserved zinc-binding HExxH motif, characteristic of zinc proteases (Montecucco and Schiavo 1993)(Fig 2.2c). The genbank accession numbers for the sequences in Fig 2.2c are A: 20137335; B: 399134; C: 1217587; D: 115188; E: 399136; F: 399137; G: 2499920. The light chain is a zinc-dependent endopeptidase, specific for the SNARE components (section 2.3) involved in neurotransmitter release (Schiavo 2000; Turton et al. 2002). Catalysis of proteolytic cleavage of SNAREs is mediated by BoNTs LC domain leaving SNARE complexes non-functional, thus, blocking the release of acetylcholine at the nerve terminals (Humeau 2000; Schiavo 2000).

```
A   DPALTLMHELIHSLHGLYGA
B   DPAISLAHELIHALHGLYGA
C   DPAVTLAHELIHAEHRLYGI
D   DPALILMHELIHVLHGLYGI
E   DPALTLMHELIHVLHGLYGI
F   DPILILMHELNHTMHNLYGI
G   DPVIALMHELTHSLHQLYGI
    ** . * *** *  * ***
```

Fig 2.2c Conserved zinc binding motif HExxH observed in all seven serotypes.

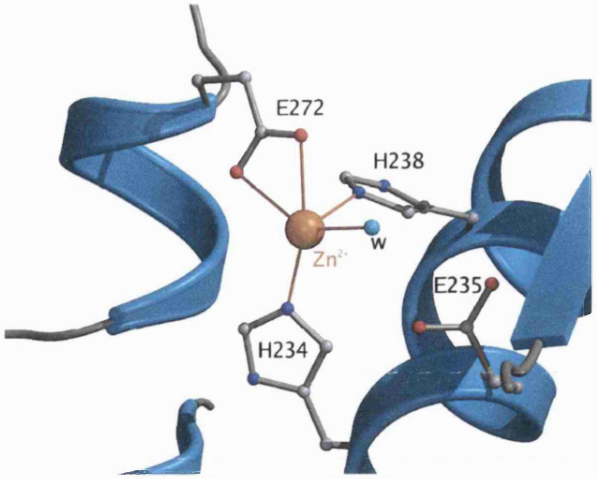


Fig 2.2d Catalytic centre of the endopeptidase. Zinc ion – *gold*; water – cyan; coordinating residues are shown as ball-and-stick model (Thiyagarajan et al., unpublished results).

2.3 Control of Neurotransmission by SNARE Complexes

Neurotransmission is controlled and regulated by a special group of proteins called SNARE (soluble N-sensitive factor attachment protein receptor), which includes VAMP (vesicle associated membrane protein), the plasma membrane protein syntaxin, and the SNAP-25 (synaptosomal-associated protein) (Bajjalieh 2001). The SNARE complex formation involves one VAMP/syntaxin or two SNAP-25 molecules, which is the requirement for vesicle fusion at the nerve terminals (Chen 2001). A mechanism for SNARE complex formation is proposed by Chen et al. (2001) (Fig 2.4).

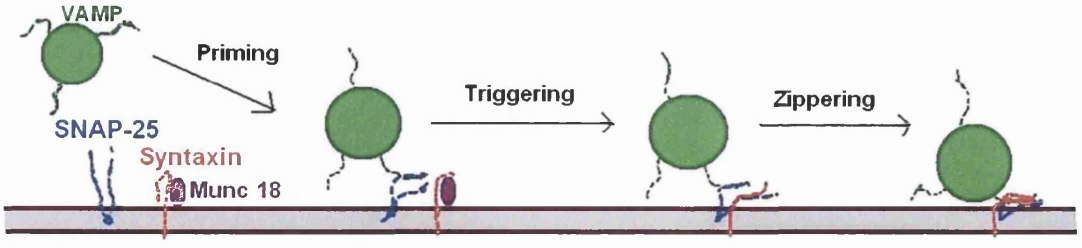


Fig 2.4 The SNARE complex formation (Chen 2001). Before priming all SNARE proteins are free in the cytosol, syntaxin is thought to be attached to

munc 18 (Haynes et al. 1999; Rickman and Davletov 2005). ATP-dependent priming binds the VAMP to SNAP-25 and Ca^{2+} dependent triggering mediates the binding of syntaxin to the VAMP - SNAP-25 complex. This loose complex 'zippers' into a tight, SDS-resistant four-helix bundle, which helps in fusion of the SNARE to the plasma membrane.

The reduction of disulphide bond between light chain and heavy chains of CNT is critical for toxins activity in the host, which is supposed to take place in the endosome. Following the separation of light chain and heavy chain the catalytic part of the enzyme is released into the cytosol. It is not yet completely understood, the mechanism about the translocation of the light chain into the cytosol. The hypothesis is that the translocation domain and the catalytic domain undergo a major structural and conformational change at endosomal pH (5.0) (Schmid et al. 1993; Kalandakanond and Coffield 2001), allowing the translocation domain to form pores on the cell membrane. The catalytic domain also undergoes a little conformational change at endosomal pH retaining its zinc site. This conformational change of catalytic domain is completely reversible at cytosolic pH (Puhar et al. 2004). Similar type of toxification is observed in case of diptheria toxin (Kagan et al. 1981).

2.4 Substrate Specificity and Toxicity

All CNTs are zinc proteases and are toxic to neuronal cells by inhibiting the release of neurotransmitter across the nerve terminals. Toxification by BoNTs or TeNTs involves the toxin binding to ganglioside receptors on the cell surface followed by intoxication – receptor mediated endocytosis (Black and Dolly 1986a; b; Montecucco 1986). The target for TeNTs and BoNTs are different. TeNTs affect the central synapses while the BoNTs affect the cholinergic nerve terminals. Studies have shown that the potential toxic fragment in the process of inhibition of exocytosis is the light chain of BoNTs and TeNTs (Ahnert-Hilger et al. 1989). The special features of CNTs are that

the toxins are synthesised as an inactive precursors, which becomes active on subsequent proteolytic cleavage in the host (Montecucco and Schiavo 1993). The presence of zinc ion is very important for these endopeptidases. Removal of zinc completely abolishes the substrate specificity and activity of CNTs in the host (Tonello et al. 1997). Substitution of other divalent metal ions retains some activity of the enzyme (Tonello et al. 1997; Eswaramoorthy et al. 2004).

Despite high sequence similarity between BoNT serotypes their substrate specificity is very selective (Schiavo et al. 1993; Binz et al. 1994; Vaidyanathan et al. 1999) (Table 2.1). Whilst the substrate specificity, BoNTs show a requirement for length of the substrate it binds (Shone et al. 1993; Pellizzari 1996). This indicates that the substrate specificity of BoNTs is not decided by the active site, but a few exosites located around the active site (Breidenbach and Brunger 2004). In turn the SNARE proteins share a common nine-residue motif, thought to undergo structural conformational change upon binding to the BoNTs (Rosseto et al. 1994; Wictome et al. 1996; Breidenbach and Brunger 2004).

Table 2.1 Substrate specificity and scissile bond selectivity of light chain of BoNTs. SNAP-25: synaptosomal associated protein, a 25 kDa protein; VAMP-2: vesicle associated membrane protein, also called as synaptobrevin II.

BoNTs	***SNARE Substrate***	***Scissile bond***
A	SNAP-25	Q197 – R198
B	Synaptobrevin II	Q76 – F77
C	Syntaxin, SNAP-25	K253 – A254, R198 – A199
D	Synaptobrevin II	K59 – L60
E	SNAP-25	R180 – I181
F	Synaptobrevin II	Q58 – K59
G	Synaptobrevin II	A81 – A82

2.5 Mechanism of action

Proposed mechanism of activation of the catalytic domain (LC) of botulinum neurotoxin (Fig 2.5):

1. Binding of the BoNT to the gangliosides (Simpson and Rapport 1971b; a) of the neuronal cell membrane. The residues involved in putative ganglioside binding include Ser1287, Trp1289, Tyr1290 and Gly1300 and is conserved in both TeNT and BoNT. The residue Asp1222 in TeNT is replaced with glutamic acid in BoNTs (Ginalski et al. 2000).

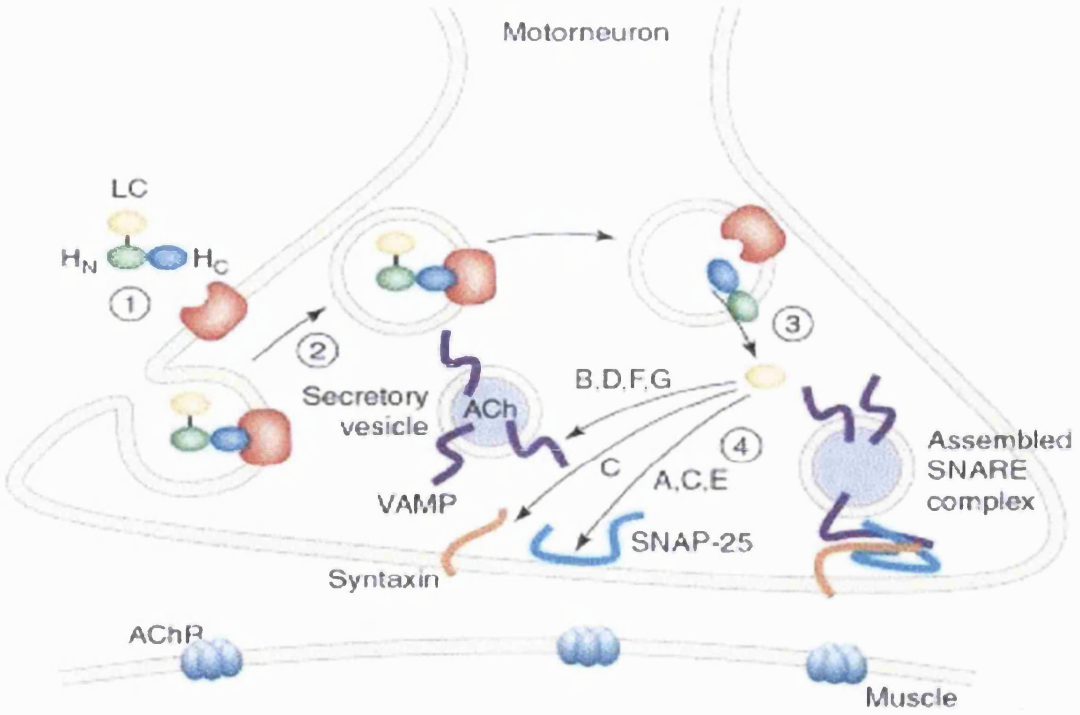
2. The bound toxin is endocytosed into the presynaptic vesicle (Montecucco 1986; Ginalski et al. 2000; Humeau 2000)

3. Following the endocytosis it is proposed that there is a pH-dependent structural change of the toxin occurs, which release catalytic domain (LC) into the cytosol (Montecucco 1994; Maksymowych and Simpson 1998; Puhar et al. 2004).

4. Inhibition of neurotransmitter release is initiated by binding and proteolytic cleavage of SNAREs.

2.6 Therapeutic use of Botulinum Neurotoxins

Botulinum neurotoxins are the causative agents of botulism, a neuroparalytic illness for which there is no drug/vaccine widely available to date. This could be a reason why this toxin is so extensively studied. But, this is not true, as the botulinum neurotoxins has much more interesting properties and used in many clinical applications. The longevity effect of the toxin makes it a suitable target in therapeutics. BoNT/A is very well known by the trade name BOTOX® and Dysport™ for its use in cosmetics to reduce wrinkles in the skin.



Turton *et al* 2002

Fig 2.5 Proposed mechanism of BoNTs entry and release of the catalytic domain into the cytosol (Montecucco 1994; Turton et al. 2002).

Other clinical applications are treatment of cervical dystonia (Greene et al. 1990), achalasia (Annese et al. 1998), blepharospasm (Jankovic and Orman 1987), spasticity (Snow et al. 1990; Simpson et al. 1996), autonomic disorders such as anal fissure (Maria et al. 1998; Naumann et al. 1999). It is studied that BoNT/A can inhibit overactive nonmotor and motor peripheral acetylcholine neurons (Silberstein et al. 2000) resulting in reduced hyperactivity of muscles and spasm. The side effects and proper analysis of drug administration indicated that BoNT/A could be used in pain relief (Silberstein et al. 2000) such as chronic pain (Rollnik et al. 2001), cervical dystonia, and migraine (Silberstein et al. 2000). Intravenously injected clostridial spores become exclusively localised to the hypoxic/necrotic tissue common to solid tumours and thus the clostridial spores could be administered as a drug delivery vehicle in the treatment of cancer chemotherapy (Minton et al. 1995).

Inhibitors of Botulinum Neurotoxins

Botulism can be spread air-borne by the spores of botulinum and hence, there is a threat that *Clostridium botulinum* could be used as a bio-weapon (Leggiadro 2000). Symptoms appear from 12 hours to 6 weeks from toxin infection and diagnosis of botulism is not a routine practice in clinical laboratories. Since there is no effective cure available yet, side-effects from clinical use of the toxin and threat as a bio-weapon makes it a hot molecule in research. Several agents to detoxify the toxicity of the toxin have been identified (Mikhailov and Mikhailov 1970), but, this could only be used to delay the onset of paralysis and not a preventive one (Simpson 1986; Deshpande et al. 1997). Several serotype specific small molecule inhibitors and peptide inhibitors of botulinum neurotoxins have been reported (Satoh et al. 2001; Anne et al. 2003a; Anne et al. 2003b; Burnett et al. 2003) but in early stages of identifying the lead compound(s) and not recommended for use in therapeutics. Further studies on toxins structural and chemical property i.e., the mechanism of toxins binding to the pre-synaptic membrane, release into the cytosol from the endocytosed synaptic vesicle and substrate specificity would need to be elucidated in order to use botulinum toxin as an effective therapeutic.

2.7 Novel non-toxic LHn fragments

Inactivation of the botulinum neurotoxin by proteolytic action of trypsin on the C-terminal residues of the binding domain and the formation of channels in bilayers by the translocation domain has been reported (Shone et al. 1985; Donovan and Middlebrook 1986; Blaustein et al. 1987; Shone et al. 1987). Hence the deadliest toxin becomes non-toxic or inactivated upon removing/losing the binding domain. This information is applied in the creation of *novel non-toxic light and heavy chain (LHn) fragments of botulinum neurotoxin* ~100 kDa in molecular weight, which comprises only the light chain (LC) and the translocation domain (Hn) of the heavy chain (HC). The non-

toxic LHn fragments are engineered to resemble the native toxin lacking the cell surface-binding domain. Application of recombinant technology in the creation of novel non-toxic fragments of botulinum neurotoxin have been reported (Chaddock et al. 2002; Duggan et al. 2002; Chaddock et al. 2004; Turton et al. 2004; Sutton et al. 2005; Foster et al. 2006). Suitable ligands are selected, designed and/or researched for use in retargeting the non-toxic LHn fragments as a therapeutic. Studies on retargeting the LHn fragments in mice and their effects have been reported (Chaddock et al. 2000a; Chaddock et al. 2000b; Chaddock et al. 2004; Sutton et al. 2005; Foster et al. 2006).

Though whole and part (catalytic domain only) of botulinum toxins structural information is available the innate properties - mechanism of toxin's neurotoxicity, substrate specificity and substrate length selectivity are not yet structurally characterised. In contrast there is no structure available for non-toxic LHn fragments. The LHn fragments retain the intrinsic neurotoxic property even in the absence of cell-binding domain/ligand, i.e., the toxin remains inactive unless endocytosed after binding to cell-surface with its binding domain/ligand (absent in LHn fragments) and activated by reduction of disulphide bond bridging between light chain and heavy chain fragments. Understanding the structure of LHn fragments and using it for structure-based drug design would help in harnessing the use of LHn molecules as a therapeutic and drug for cure in botulinum infection. Initial homology modelling suggested that the LHn fragments should also possess similar structure to the native toxin, the only difference being is the absence of the binding domain (Fig 2.6). Yet this is only a model and a proper structural and chemical characterisation is a prerequisite to understand these novel molecules to be used in therapeutics and in drug design.

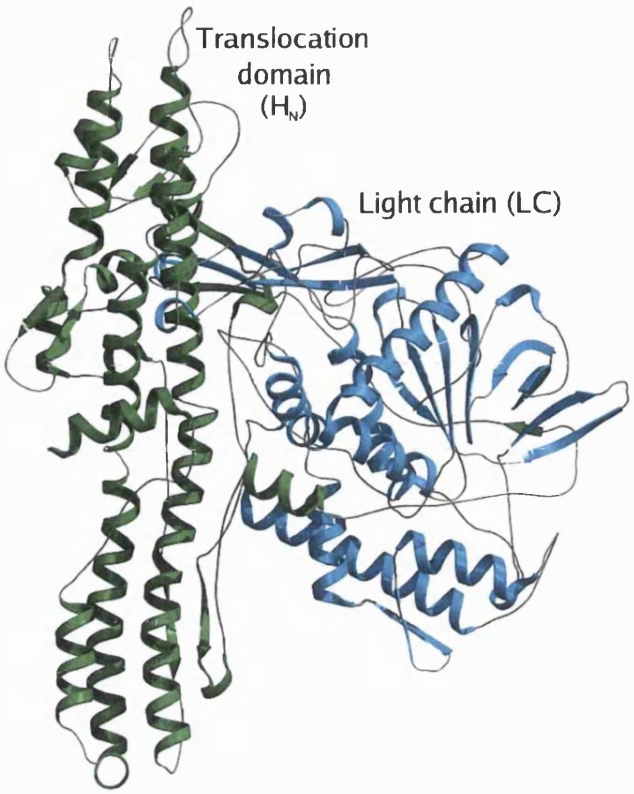


Fig 2.6 Ribbon diagram representation of homology model of LHn/A (Details of homology modelling in section 2.8)

2.7.1 Initial crystallisation screens for LHn/A and LHn/C

Initial crystallisation screens of LHn fragments were performed with the recombinant proteins provided by Dr. John Chaddock. The LHn proteins were in buffer 50mM Hepes pH 7.2 and traces of phosphate salts. The proteins were provided in the lyophilised form. Initial protein concentration of ~10 mg/ml was used for standard screens (Structure screen I and II (Molecular Dimensions, Ltd, UK), PEG/Ion screen, and Ammonium sulphate grid screen (Hampton Research, USA)). The crystallisation setups were carried using the hanging drop vapour diffusion technique using 24-well linbrow plates (Molecular Dimensions Ltd., UK). The drops containing equiproportion of protein to mother liquor is equilibrated against 500-1000 µl of the reservoir solution and incubated at 16° C. Immediate precipitation was observed in

most of the structure screen conditions. Drops were observed for any hit after 24 hours of setup. Repeated checks for months, and self-designed screens (Combinations of PEG, Ammonium Sulphate, MPD, different buffers at different pH, additives, use of detergents) and some of the conditions previously reported for the whole or part (light chain) of the toxin were also screened. Whilst waiting for a hit from the crystallisation setups, report of successful crystallisation and improvement in resolution of the diffraction data by surface engineering of protein molecules by Derewenda *et al.* directed to try a different approach of mutating the surface residues or residues that might hinder the crystal contacts of LHn fragments (Derewanda et al., 2004).

2.7.2 Site-directed Mutagenesis to Aid Crystallisation of LHn Fragments Type A

All the work relating cloning, mutation, expression and purification were carried out at Health Protection Agency (part of this organisation is now called Syntaxin Ltd.), Porton Down, Salisbury. The clones for LHA-S877, LHC-EGF, and LHC-HT were provided Dr. John Chaddock to start the experiment except where mutation or reengineering or recloning the gene into a different vector is carried out. Modified pMAL vectors were used for all cloning work and expression systems (section 2.8).

The LHn/A is purified in the native and mutant forms. LHn/C is purified either with C-terminal His-tag and/or as a fusion protein with Epidermal Growth Factor (EGF) and hereafter referred as LHC-HT and LHC-EGF respectively. All the genes were cloned and expressed in *Escherichia coli* BL21 cells either with cleavable N- or un-cleavable C- terminal His-tag and with cleavable N-terminal MBP-tag or without N-terminal MBP-tag following induction with IPTG (isopropyl-beta-D-thiogalactopyranoside). A detailed protocol from expression to purification and assays is given in section 2.8 along with the gel pictures.

The LHn/A protein initially used for crystallisation trials were 871 amino acids in length which was later increased to include the full translocation domain (877 amino acids in length). The increase in length was to stabilise the translocation domain and thought the extended version of LHn/A would favour crystallisation unlike the short version (871 amino acids). Molecular modelling of LHn/A and comparative analysis with other available whole toxin structure gave a direction to proceed. The whole length of LHn/A-S877 (an extended version of 871 amino acids) including the complete translocation domain sequence (877 amino acids in length) was submitted to the Swiss-protein modelling server, an automated protein homology modelling server.

2.7.3 Homology modelling of LHn/A

Molecular modelling of LHn/A is done using the Swiss-MODEL (http://swissmodel.expasy.org//SWISS-MODEL.html), an automated comparative protein modelling server. From the sequence comparison results the structural data of botulinum neurotoxin A (PDB: 3BTA) (Lacy et al., 1998) was selected as the target template for homology modelling and for further comparative studies. The LHn/A model were completely biased by its template. Hence, a round of energy minimisation was carried using the software suite Crystallographic and NMR System (CNS) (Brünger 1998). The side of the translocation domain interacting with binding domain in the whole toxin is solvent exposed in the non-toxic LHn fragments. Hence the suspection is on the crystal contacts of the solvent exposed part of the translocation domain of LHn fragments. Hence to aid crystallogenesis site-directed mutagenesis experimental studies were done.

All the mutation works were carried out in the LHn/A-S877 (hereafter referred as LHA-S877) background. Three strategies were tried to improve the stability of the protein regions in the translocation domain leaving the catalytic domain undisturbed. LHA-S877 is also expressed, purified and screened for crystallisation trials.

- Stabilisation of loop 1 by mutating K625A and I626A. (Fig 2.7a). The residues K625 and I626 in the loop were thought to be involved in crystal contact of LHn/A. Based on interactions observed for these residues with the binding domain in whole toxin, predictions were made that the residues K625 and I626 could disturb the crystal formation.
- Stabilisation of helices 2 and 4 (Fig 2.7b) by introducing a disulphide bridge between the residues V793 and T865. These two mutations were considered to lock the helix 4 to helix 2 by forming a disulphide bridge.
- Mutation at the tip of helix 3 to favour the interaction with the neighbouring polar residues of helix 2 (Fig 2.7c).

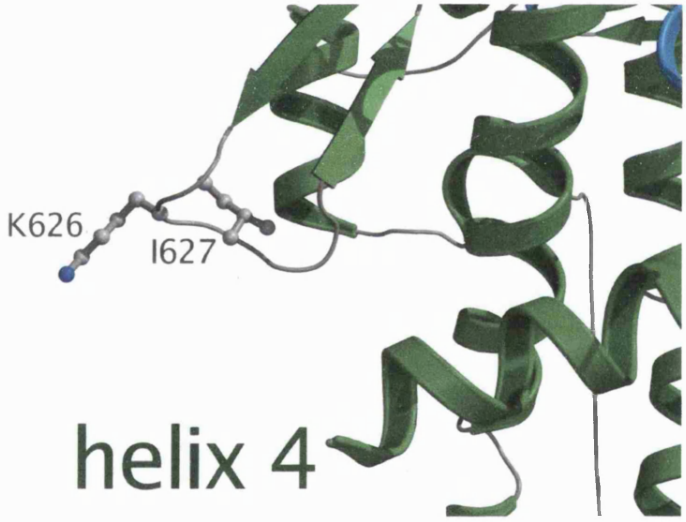


Fig 2.7a Homology model of LHA-S877 - Residues in loop K625 and I626 are shown as ball-and-stick model – grey. These two residues were mutated to alanine.

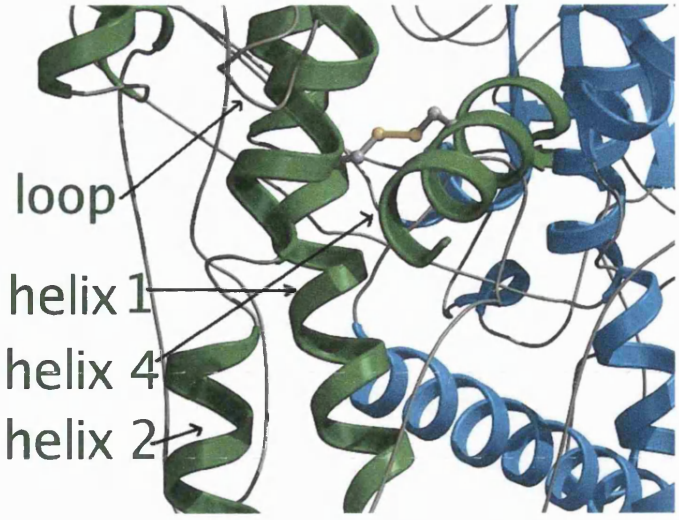


Fig 2.7b Homology model of LHA-S877 - Residues between helices 2 and 4, V793 and T865 are mutated to cysteines. Disulphide bridge is shown as yellow stick, the cysteine residues in grey ball-and-stick model.

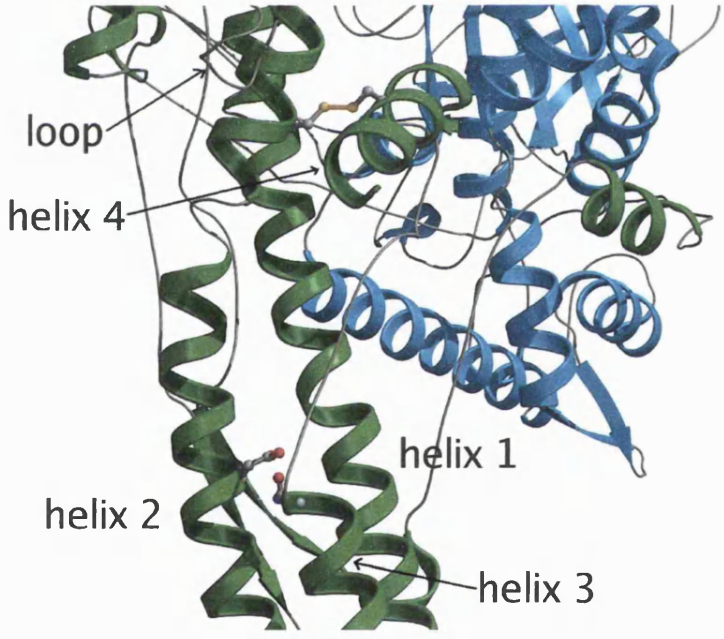


Fig 2.7c Homology model of LHA-S877 - Residue in the loop region between helix 2 and 3, L845 is mutated to either lysine or asparagine.

2.7.4 Preliminary Diffraction Data for LHA-S877, LHC-HT and LHC-EGF Fusion Protein

2.7.4.1 Materials and Methods

LHn fragments of type A and its mutants, and type C with N-terminal MBP-tag and a C-terminal His-tag and as a fusion protein with EGF were expressed in *E.coli* BL21 cells in Terrific Broth, which is inoculated from a overnight culture. The cultures were grown until an optimal optical density at 600 nm is reached prior to induction with IPTG and further grown for ~20 hours (section 2.8). The cells were harvested by centrifuging the cultures at ~4000 rpm (revolutions per minute) for ~20 min and the pellets were resuspended in the lysis buffer and stored at –80°C. Cells were lysed by sonication and spun at 18000 rpm in a Beckman JLA rotor for ~25 min. The protein was expressed in soluble form, hence the pellet is discarded and clear lysate was loaded on to the Ni^{2+}-chelating sepharose column. The bound proteins were eluted with different concentration of imidazole and ran on an SDS-PAGE gel to identify the fractions containing the protein of interest. To ensure that the maximum purity is achieved the eluted sample was again ran on second Ni^{2+}-chelating sepharose column, this time the His-tag or MBP tag in the N-terminal of the LHn fragments are cleaved by Factor Xa treatment (section 2.8). Final tuning in purification step was achieved by running the purified sample through size-exclusion chromatography column (section 2.8). The sample fractions were run on SDS-PAGE gels and checked for purity (Fig 2.15). The fractions were pooled and concentrated for crystallisation trials. The concentration of the protein samples was either measured using ultraviolet absorbance spectrum or BCA assay. The concentrated sample was finally ran on SDS-PAGE to check the final purity of the sample and stored at –20°C but not lyophilised.

Crystallisation of LHn fragments were carried out using the standard crystallisation matrix bought from Molecular Dimensions Ltd., UK, and Hampton Research, UK. Self-designed screens were also screened. All the

crystallisation setups were performed using the hanging drop vapour diffusion method as described in section 2.7.1.

2.7.4.2 Crystals of LHA-S877, LHC-EGF and LHC-HT

LHA-S877

Crystals of LHA-S877 were observed in two of the structure screen conditions after a period of ~10 months. Protein concentration used was 10 mg/ml. The conditions were

1. 30 % PEG 4000; 0.2 M Ammonium Sulphate; 0.1 M Tris-HCl pH 8.0
2. 0.8 M Ammonium Sulphate; 0.1 M Sodium citrate pH 5.0

Crystals obtained from condition II was not eproducible but the crystals diffracted to 6.0 Å resolution (Fig 2.8a). The diffraction pattern was measured at SRS, Daresbury at PX station 14.1 (λ = 1.488 Å) with crystal oscillation range of 1°, 120 sec exposure to X-radiation at cryo temperature of 100K. The crystal was cryo-protected with 25 % glycerol prior to mounting on the goniometer. There is no crystal picture taken for this condition. In contrast to condition 2, condition 1 was eproducible and used for further optimisation trials. Hundreds of tiny crystals (Fig 2.8b) grow in a period of 7-10 days using condition I. These crystals were not suitable for diffraction studies. Hence, optimisation of the condition was carried out using additives such as glucose, MPD (2-methyl-2,4-pentanediol), glycerol, metal ions like nickel, copper, iron were also used. Nothing appeared in most of the additive screens for weeks. Detergents were also used as additives bearing in mind that the LHn fragments have a transmembrane region, the Hn domain that might help in growing better crystals. One of the conditions using Hecameg detergent gave long needle crystals (Fig 2.8c). Optimisation of this condition is currently underway.

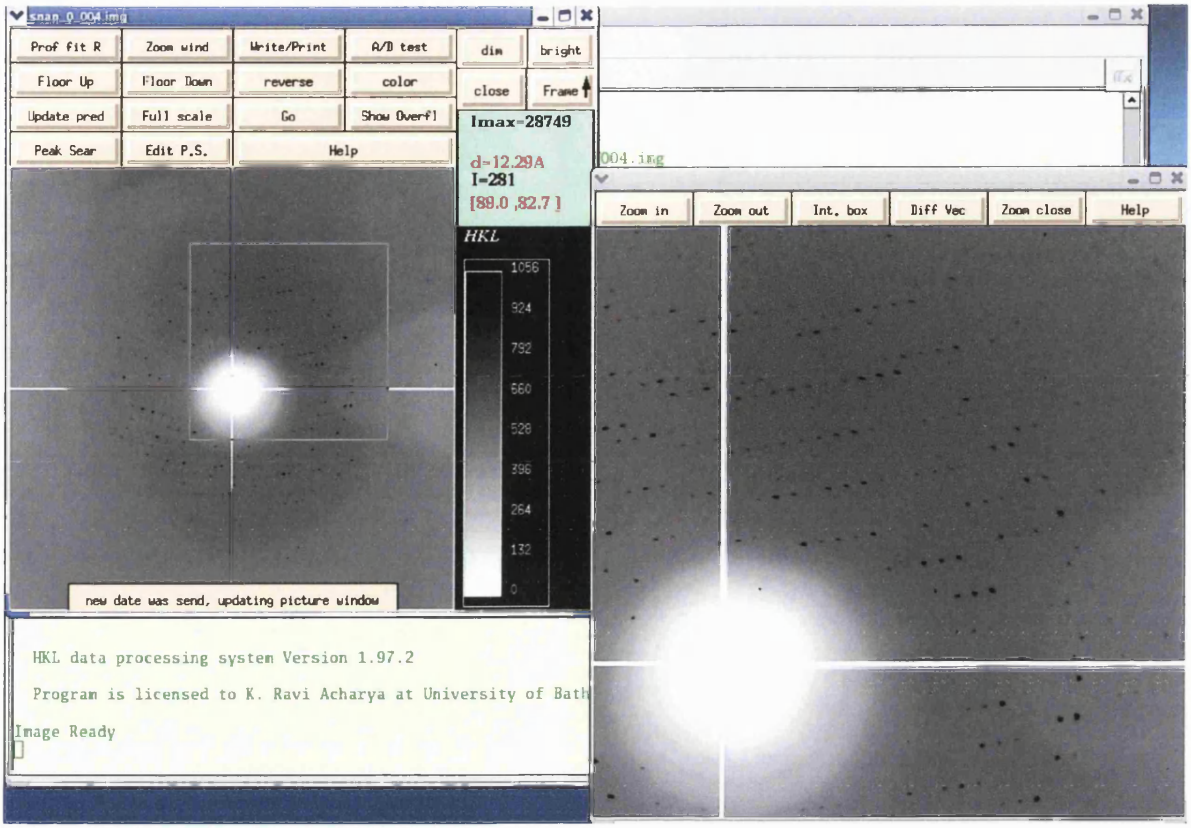


Fig 2.8a First X-ray diffraction pattern of LHA-S877. Diffraction image recorded at PX station 14.1 (λ = 1.488 Å), SRS, Daresbury with oscillation range of 1°, exposure = 120 sec.

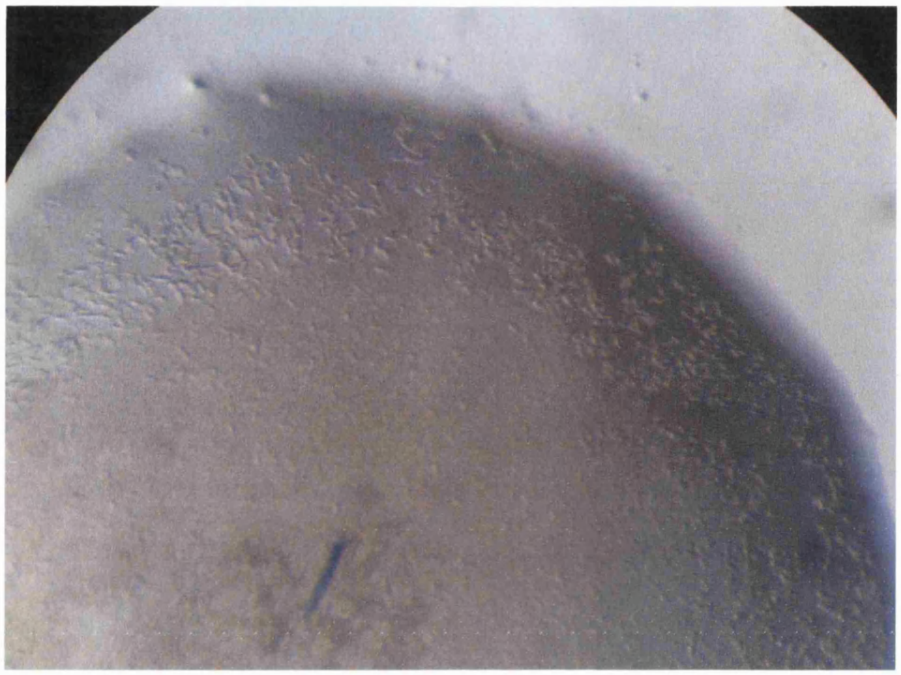


Fig 2.8b Microcrystals of LHA-S877. 30 % PEG 4000; 0.2 M Ammonium Sulphate; 0.1 M Tris-HCl pH 8.0

Fig 2.8c Needle crystals of LHA-S877 using hecameg detergent as additive.

LHC-EGF

Crystals of LHC-EGF were observed in one of the structure screen conditions. The crystals appeared after a month from setup date. Protein concentration used was ~8 mg/ml. The condition is

10 % PEG 1000 and 10 % PEG 8000

Thin plate like crystals (Fig 2.9a) grew in repeating the experiment with same condition. Optimisation with varying concentrations of the precipitant or addition of additives and detergent screen did not give any better result. The crystal did diffract to ~4.0 Å resolution. The diffraction pattern was measured at SRS, Daresbury at PX station 14.2 (λ = 0.979 Å) with crystal oscillation range of 1°, 120 sec exposure to X-radiation at cryo temperature of 100K. The crystal was cryo-protected with 25 % glycerol prior to mounting on the goniometer. The crystal lost its diffracting strength after a couple of snapshots (Fig 2.9b).

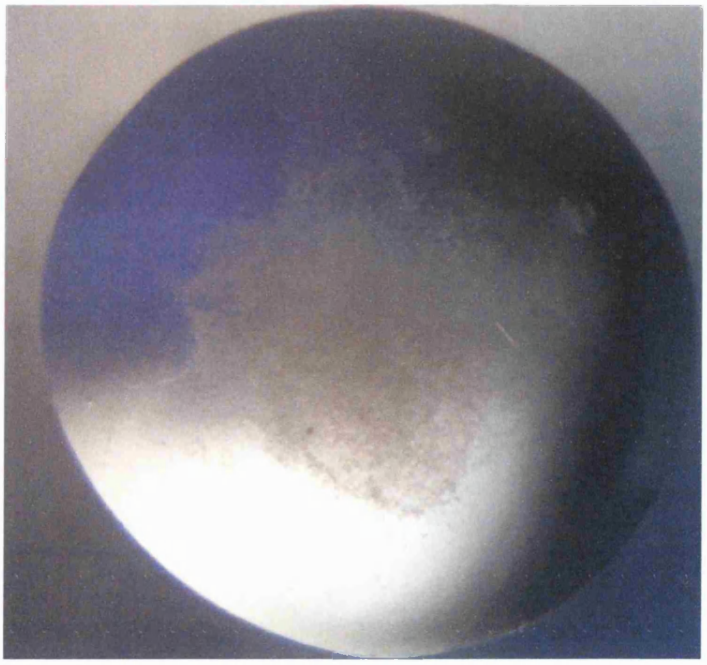


Fig 2.9a Thin plate crystals of LHC-EGF

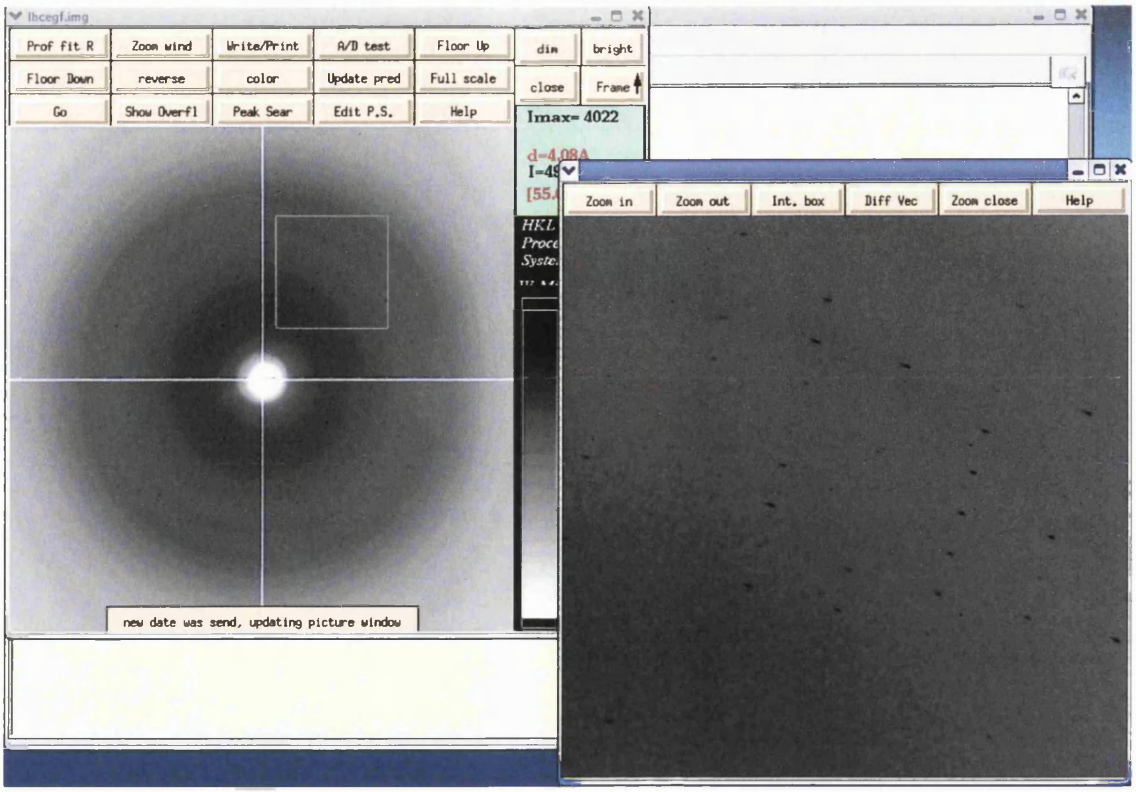


Fig 2.9b First X-ray diffraction pattern of LHC-EGF. Diffraction image recorded at PX station 14.2 (λ = 0.979 Å), SRS, Daresbury with oscillation range of 1°, exposure = 120 sec.

LHC-HT

Crystalloids of LHC with C-terminal His-tag (Fig 2.9c,d) were observed in two of the structure screen conditions. Protein concentration used was ~2 mg/ml. The crystals appeared in less than a week from setup date. The conditions are

1. 0.1 M Sodium citrate pH 5.6; 1.0 M Ammonium dihydrogen phosphate
2. 4.0 M Sodium formate

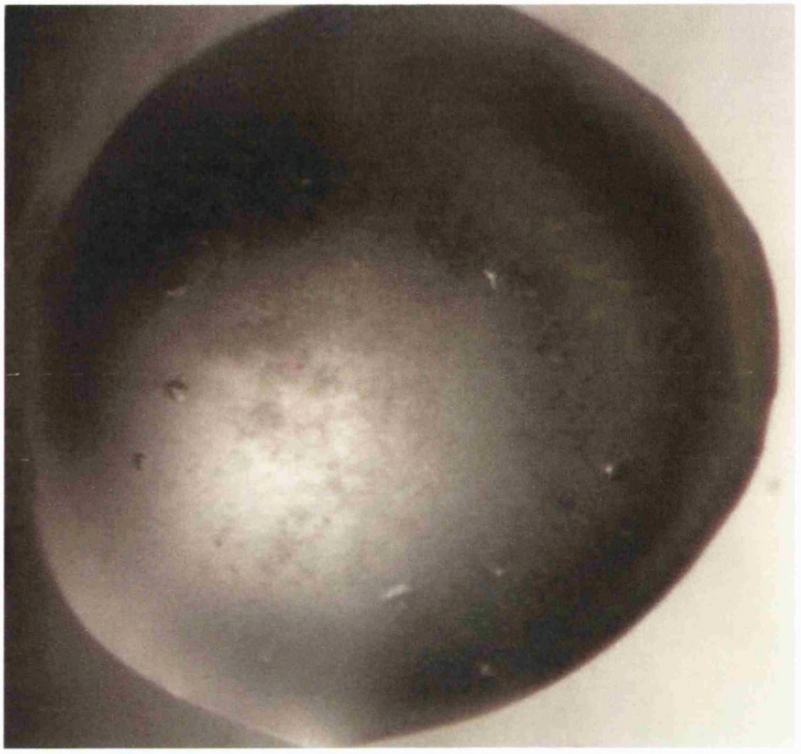


Fig 2.9c Condition 1: 0.1 M Sodium citrate pH 5.6; 1.0 M Ammonium dihydrogen phosphate

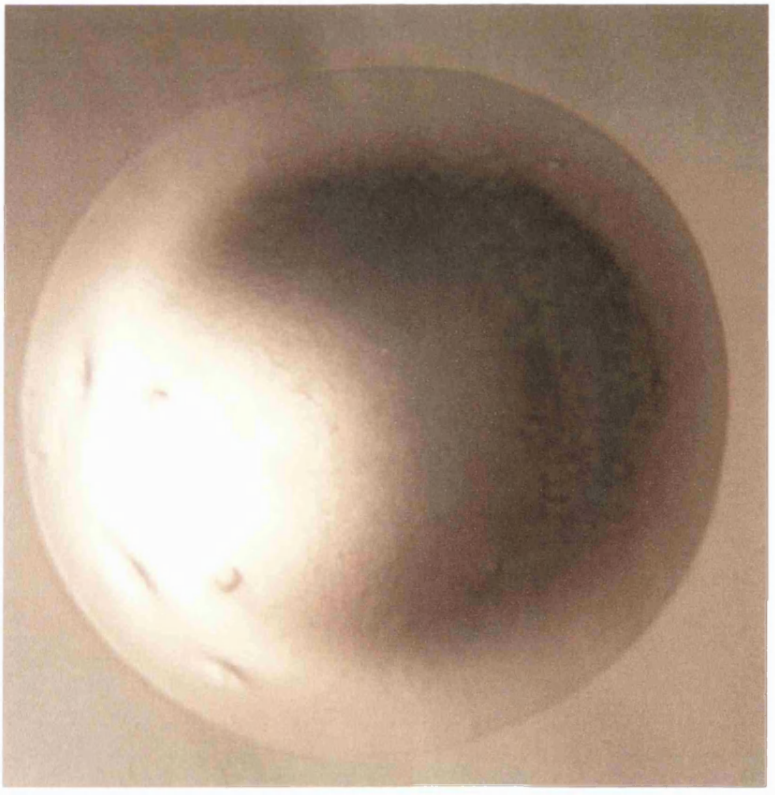


Fig 2.9d Condition 2: 4.0 M Sodium formate

2.8 Cloning, Expression, and Protein Purification of recombinant LHn Fragments

Following the homology modelling and analysis of the recLHn/A fragment model, following interpretations were made: the loop region interacting (section 2.7) with the binding domain H_C (absent in *rec*LHn) might interfere with crystal packing (due to high thermal vibrations) and thus not allowing the protein to form crystals. These observations lead us to try some mutational studies on this region. Six residues were K625, I626, D628, K646/Y647, D650 and I656, which might interfere with the crystal packing, and also to extend the H_N domain by a few residues (to restrict the movement of the loop). The desired mutations were K625A, K625A+I626A, I626A, triple mutants (V793C, T865C, L845K/N), and LHn/A truncation at residue I830.

The mutations were introduced in the S877 extension (S877 - six residues were added to the C-terminal (terminating with residue S) of the H_N domain which was previously 871 in length). Since all the mutation and/or truncation work were only in H_N domain of LHn/A the pTO vector with H_N domain only used for PCR reactions unless otherwise mentioned. The LHn/A is cloned in an modified pMAL vector (Fig 2.10) with restriction sites Pst I and Xba I for H_N domain and the whole length of the LHn domains of serotype A and/or C were between restriction sites BamH I and Hind III. All the mutations were carried out in the H_N extension (S877) template (synthesized with Pst I and Xba I restriction sites) in the pTO vector.

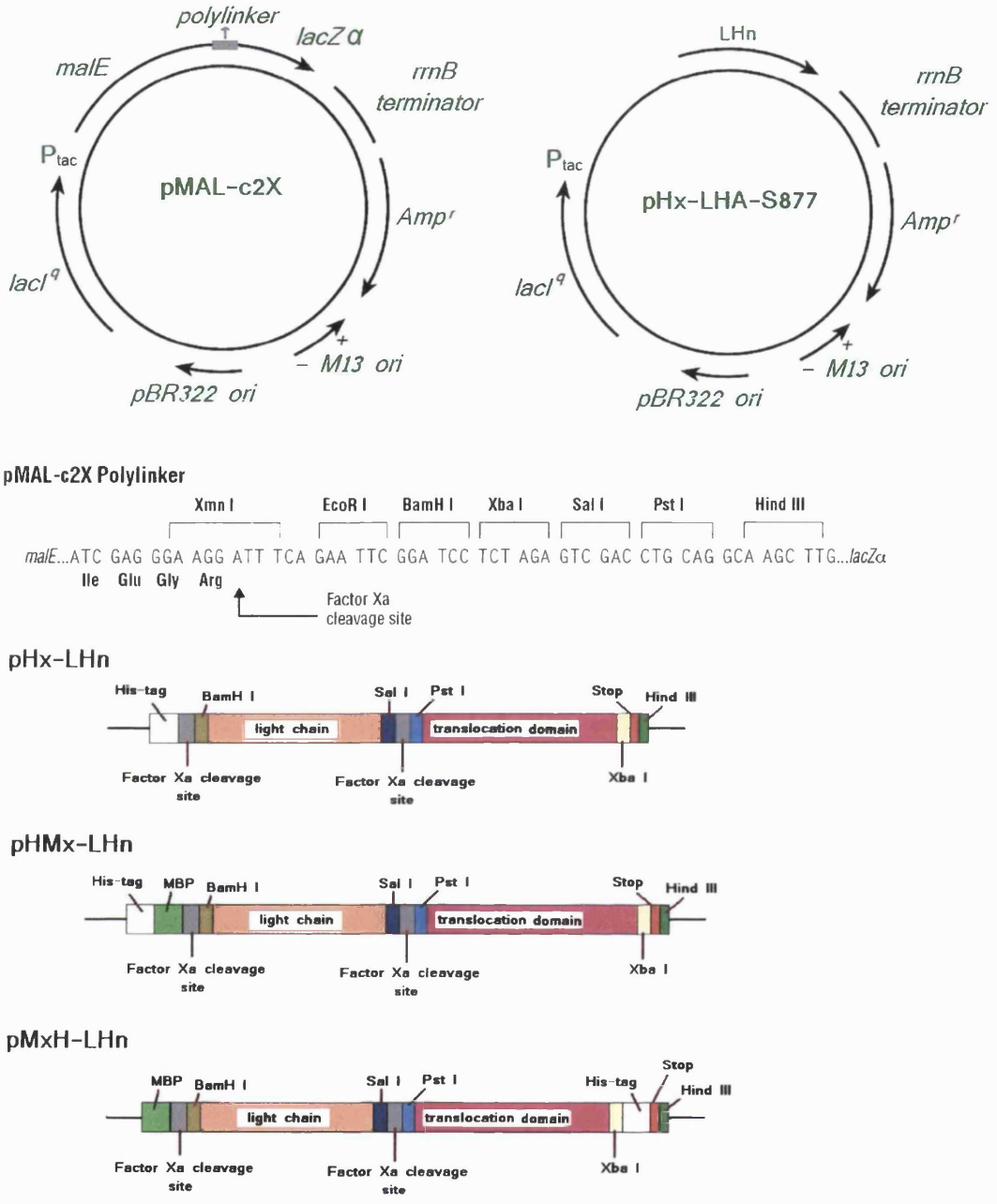


Fig 2.10 A simple diagramatic representation of the plasmid constructs

2.8.1 Materials and Methods

Mutant *rec*LHn expression and purification was done at HPA, UK. The BL21 chemically competent cells bought from Novagen (catalogue number 69499-4), TOP10 chemically competent cells bought from Invitrogen (catalogue number C4040-06). pTO vector is the pCR 4-TOPO vector from Invitrogen. pHx vector was the one prepared in Dr. John Chaddock's laboratory, Health Protection Agency (part of this organisation is now called Syntaxin Ltd.), Porton Down, Salisbury. The His-tag was inserted into the pMAL-c2x vector from New England Biolabs (NEB) in place of the MBP tag. A small section of the pMAL plasmid has been removed to make it non-transmissible. The mutant primers were PCR amplified, sequence verified, cloned into expression vectors and expressed. All the primers for mutation work were synthetically prepared and bought from MWG and/or Sigma-genosys, UK. All the reagents for PCR reaction were bought from Stratagene, UK. The DNA purification reagents were bought from Qiagen, UK, Promega, UK and Stratagene, UK. Chemical compounds and salts used for protein expression and purification were bought from Sigma, UK. Readymade SDS-PAGE gels were bought from Invitrogen, UK.

Expression of *rec*LHn fragments:

NOTE: All recLHn plasmids used were ampicillin resistant. Following constructs of LHn fragments were PCR amplified and/or cloned, expressed and purified.

1. pHx-LHA-S877
2. pHx-LHA-K625A
3. pHx-LHA-I626A
4. pHx-LHA-K625A-I626A
5. pHMx-LHA-S877
6. pHMx-LHA-I830 (Truncated LHn/A)
7. pMxH-LCC

8. pHMx-LHC-His-Tag (pHMx-LHC-HT)
9. pMxH-LHC-EGF
10. pHMx-LHA-CKC-S877 (V793C/L845K/T865C)
11. pHMx-LHA-CNC-S877 (V793C/L845N/T865C)

Description of vectors:

pHx- = cleavable N-terminal His-tag

pHMx = cleavable N-terminal His-MBP-tag

pMxH = cleavable N-terminal MBP-tag and non-cleavable C terminal His-tag

Molecular weight of LHn fragments:

LHA-S877	~100 kDa
LCC	~50 kDa
LHC-EGF	~110 kDa

PCR amplification of the mutants:

Each PCR reaction (for single mutation or truncation work) were carried out in a PCR machine using the standard protocol recommended by Stratagene and/or Qiagen. The T_m values for primers were calculated theoretically using the formula

$$T_m = 81.5 + 0.41\ (\%GC) - 675/N - \%mismatch$$

where, T_m - melting temperature

%GC - percentage of GC content of primer

Designed primers for constructs 2, 3, 4, 6, 7/8, 10 and 11 were commercially synthesised and bought from MWG and/or Sigma-genosys. PCR products were sequence verified from Lark Technologies Inc.

A general PCR protocol used is given below:

Materials required: 10x PFu buffer, plasmid DNA template, primer mixes, dNTPs, DNA polymerase.

For a 50 µl reaction: add 5 µl of PFu buffer, 1 µl of template, 2.5 µl of primer mix at ratio 1:1, 1 µl of dNTP, 39.5 µl of double distilled water and 1 µl of Turbo DNA polymerase.

A three-step PCR cycle is given below:

1. Initial activation step at 95° C for 1 min
2. Repeat I-III for 16 cycles
    I. Denaturation step at 95° C for 1min
    II. Annealing step at 55° C for 1 min
    III. Extension step at 68° C for 12 min
3. Final hold at 4° C

1 µl of Dpn I enzyme at a concentration of 10 units/µl was added to the final PCR product and incubated at 37° C for approximately 3 - 5 hours. The Dpn I treated PCR products were ligated into pTO vector and then transformed into TOP10 competent cells. The transformed products were applied on an ampicillin resistant agar medium and incubated at 37° C overnight for colonial growth of the plasmid.

All PCR reactions were not straightforward. In case of construct number 6 for pHMx-LHA-I830 truncation variable concentration of $MgCl_2$ was used in a temperature gradient PCR with different annealing temperature from 45° C – 65° C (Table 2.4) and a different DNA polymerase (HotStarTaq) from Qiagen were also tried.

Table 2.2 Gradient PCR setup for pTO-H_N-I830

DNA polymerase	$MgCl_2$ (mM)	Annealing temperature (° C)
HotStarTaq (Qiagen)	0.0/1.5/2.5/4.5	45/55/65
PFu Turbo (Strategene)	2.5	45/55/65

The PCR reactions setup with PFu turbo in the presence of $MgCl_2$ and at different annealing temperatures was successful and sequence verified.

In case of triple mutants there was a serious problem with one of the mutant primers designed (V793C). Despite different temperature gradient concentration and $MgCl_2$ concentration, the mutation failed to give a successful result. Hence the primers were redesigned; one shorter and one larger than the primer used and the PCR reactions were setup. The short length primer (25 bases in length) gave a successful result by following the standard Stratagene protocol. Addition of $MgCl_2$ did not help the PCR reaction, instead gave negative results. But, temperature gradient from 45° C – 68° C did not exhibit any difference in the reaction cycles.

Ligation of the sequence verified constructs:

All the mutation and/or truncation work were carried out in pTO vector with restriction sites Xba I and Pst I. The pTO vector and pHx/pHMx/pMxH modified pMAL vectors were double digested with Xba I and Pst I at 37° C on a water bath for approximately 1 – 2 hours. The digested samples were then run on a 0.8 – 1 % agarose gel electrophoresis at 110 V - 120 V for approximately 50 minutes. The required fragments were gel purified using Promega DNA purification kit using the protocol recommended. The pTO-H_N domain with the required mutation and/or truncation is double-digested, gel purified and ligated into the modified pMAL expression vectors (Fig 2.17) using Stratagene ligation kit.

Transfer of LHA-S877 from pHx- vector to pHMx- vector were tried to see any better expression of the protein in a different plasmid with MBP at N-terminus. The pHx-LHA-S877 and pHMx- vectors were sequentially digested with BamH I and Hind III restriction enzymes and any one of the digested sample was treated with calf intestinal mucosa (CIP) to avoid recircularisation of the plasmid by removing the 5'-PO_4^- from the DNA. Care was taken not to treat both the insert and plasmid with CIP before the ligation reaction was setup.

An overview of expression of the constructs:

- Overnight culture (50 - 100 ml) of the respective *rec*LHn strain was grown in either Modified TB (Terrific Broth) / LB (Luria Broth) medium in the presence of 1 ml of 100 mg/ml ampicillin (final concentration of 0.1 mg/ml) at 37° C in a conical flask shaking at 250 rpm.
- To 1 l of Modified TB 1 ml of 100 mg/ml Ampicillin and 10 ml of 20% glucose were added and the contents were emptied into a 2 l conical flask.
- The TB medium in 2 l conical flask was inoculated with 10 ml – 25 ml of overnight culture and left in the shaking incubator (250 rpm) at 37° C.
- The absorbance/optical density at 600 nm (OD_{600nm}) were measured every hour.
- The temperature was lowered to 16° C once the optimal OD is reached (~0.5 – 0.6) and left for an hour to equilibrate.
- The cultures were induced with 1 ml of 1 M IPTG (final concentration of 1 mM) and grown for approximately 20 hours at 16° C.
- After ~20 hr the cultures were harvested, the cell paste was weighed, resuspended in a small volume of buffer (0.1 M Tris-HCl pH 8.0, 0.15 M NaCl) and storedat -80° C.
- The average amount of cell paste yield per litre of culture was ~10 mg. In case of LHA-S877 it was in the range of ~8- 10 g and LHC-EGF/LHC-HT it was ~10 – 15 g.
- 25 µl of the culture were taken to run on an SDS-PAGE system to check the expression.

- A sample gel picture is shown (Fig 2.11).

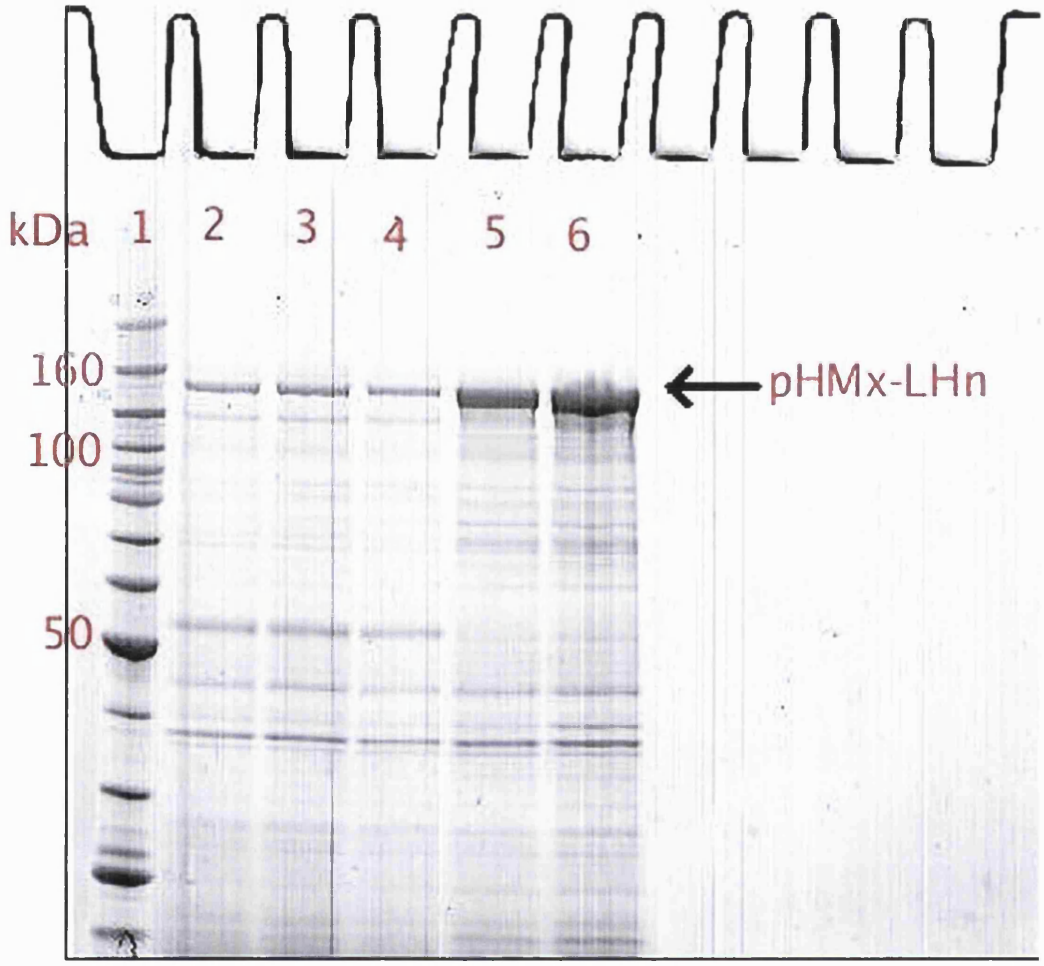


Fig 2.11 A sample expression SDS-PAGE picture for the triple mutant LHA-CKC-S877 and LHC-HT. Lane 1 is the BenchMark™ Protein Ladder (Invitrogen). Lanes 2-4 are the expression samples of pHMx-LHA-S877; lanes 5-6 are for pHMx-LHC-HT

Protein Purification of recLHn fragments :

Sonication:

The cells were disrupted using sonication method at ~20-22 micron amplitude. The cell paste was topped up with the buffer prior to sonication (say ~10-13 g of cell paste was topped up to a volume of 40 – 50 ml i.e. 1 g of cell paste in 5

ml of lysis buffer). A total of 10-12 cycles with 30 s on and 30 s off were carried. The cell paste was always kept on ice while sonicating, so as to keep the sample cool and to protect from the heat generated by sonication.

The sonicated sample was spun at 18,000 rpm at 4° C for 20 min. The clear lysate was decanted into a duran or falcon tube. The protein was expressed as soluble fraction. The equipments used for affinity chromatography were ÄKTA prime™/ÄKTA FPLC™, GE Healthcare. The columns used were XK 16 and/or XK 26, Amersham Pharmacia Biotech (now called GE Healthcare). For SDS-PAGE; NuPAGE® Novex 4-12 % Bis-Tris Gels, NuPAGE® LDS Sample Buffer, NuPAGE® MOPS SDS Running Buffer from Invitrogen were used to run the protein samples. The gels were stained using SimplyBlue™ SafeStain Microwave protocol. The SDS-PAGE scanning instrumetn used was densitometer from Molecular Dimensions, UK. The concentrators used to concentrate protein were bought from Vivascience (polyethersulfone (PES) membrane). For BCA assay measurements Fluostar Optima from BMG labtechnologies was used.

Protein purification protocol short in detail:

Choice of Tris-HCl pH 8.0 buffer for protein purification:

Initially 50 mM Hepes pH 7.4 with 0.15 M NaCl was used as buffer in LHn fragment purification. But, later 0.1 M Tris-HCl pH 8.0, 0.15 M NaCl was used because of the following reasons:

1. Protein precipitates on concentration more than 2 mg/ml in 50 mM Hepes pH 7.4, 0.15 M NaCl. Hence, the samples are lyophilised (~0.2 – 1.2 mg) in the presence of sodium phosphate buffer pH 7.4, 0.075M NaCl and 3 % Trehalose.
2. The lyophilised samples were resuspended in water or in the sample buffer for crystallisation trials (No precipitation observed after resuspension).
3. Crystallisation trials gave either precipitate or salt crystals.

4. So, decided to go for an alternate buffer with/without salt in order to improve the method of purification.
5. Using Tris-HCl pH 8.0, 0.15 M NaCl as buffer, the final concentration of the protein was achieved to 10 – 30 mg/ml, but the concentration was made to be 10 mg/ml.
6. Using 50 mM Hepes pH 7.4, 0.15 M NaCl as buffer the protein was eluted at 100 mM imidazole concentration where other impurities were also eluted. But, using 100 mM Tris-HCl pH 8.0, 0.15 M NaCl as buffer the protein was eluted at lower concentration of imidazole (30 – 40 mM) containing less impurities when compared to the elution at 100 mM imidazole. In Fig 2.19, Lanes 13-18 were eluted at 30 – 40 mM imidazole. Lane 20 was eluted at 100 mM imidazole.
7. Variable concentration (10 – 250 mM imidazole) of elution buffer was prepared using different volumes of imidazole stock solution in different volumes of buffer (0.1 M Tris-HCl pH 8.0, 0.15 M NaCl).
8. This choice of buffer helped to keep the protein sample in solution. Success has been achieved in crystallisation using Tris-HCl pH 8.0, 0.15 M NaCl as buffer.

Affinity Chromatography 1:

i. Ni-affinity chromatography was used.
ii. The proteins have either N-terminal cleavable His-Tag (*pHMx*) or non-cleavable C-terminal His-Tag (*pMxH*), which will help the proteins to chelate to Nickel ions in the column.
iii. The supernatant after sonication was loaded on to the Nickel Chelating Sepharose column, which was charged with 100 mM Nickel sulphate and equilibrated with buffer 0.1 M Tris-HCl pH 8.0, 0.15 M NaCl.
iv. Variable concentration of Imidazole (10 – 250 mM) was used to elute the protein.
v. The elution buffer used was 0.1 M Tris-HCl pH 8.0, 0.15 M NaCl and (10 - 250 mM) imidazole.

vi. After elution the fractions were run on a SDS-PAGE system to check for the right molecular weight bands.

vii. The right fractions were then transferred into a dialysis membrane and dialysed against buffer containing 0.1 M Tris-HCl and 0.15 M NaCl overnight at 4° C.

viii. Nickel from chelating sepharose column was washed with 20 mM EDTA and equilbrated and stored in 20 % ethanol. The column was recharged with 100 mM Nickel sulphate and equilibrated with buffer (0.1 M Tris-HCl pH 8.0, 0.15 M NaCl) prior to next run.

ix. A sample gel picture is shown in Fig 2.12.

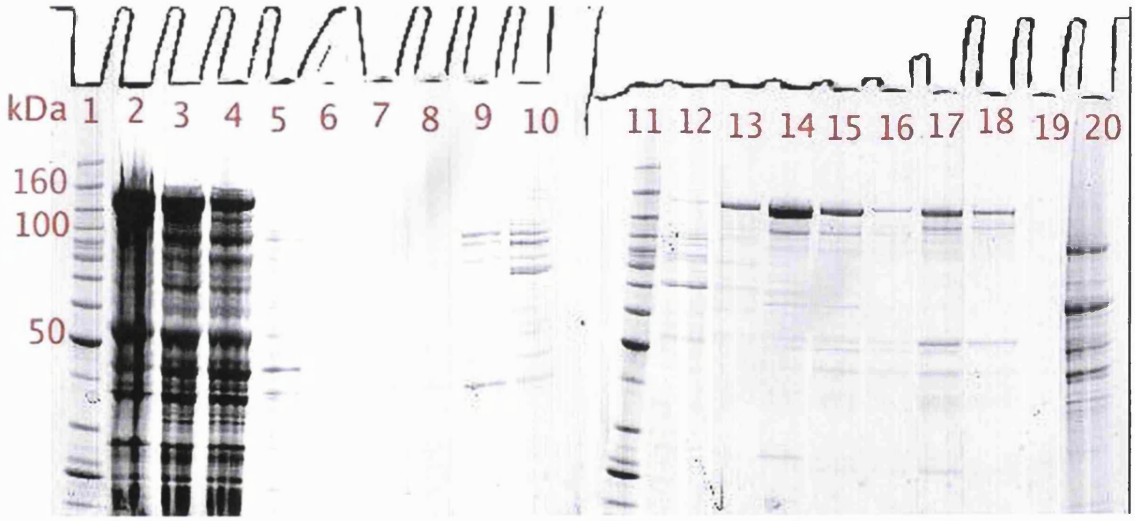


Fig 2.12 SDS-PAGE after affinity chromatography 1. Lane 1 and 11 are the BenchMark™ Protein Ladder (Invitrogen). Lane 1: crude sample (soluble) after sonication; lane 2-5: flow through from load; lanes 6-20: eluants with different concentration of imidazole. Lanes 6-8: 10 mM imidazole; Lanes 9-12: 20 mM imidazole; Lanes 13-19: 40 mM imidazole; Lane 20: 100 mM imidazole. The advantage of using higher pH buffer like Tris-HCl over Hepes buffer is that the associated impurities at 100 mM imidazole eluant is avoided by eluting the protein of interest at a low concentration of imidazole. Also LHn proteins were very stable at high concentration and remains in solution at this pH.

Factor Xa Treatment :

x. Next day the dialysed sample was collected in a duran or a falcon tube and UV absorbance at 280 nm was measured. The reference sample used was the dialysis buffer.

xi. Based on the difference in absorbance at 310 nm and 274 nm the protein concentration was calculated using the formula "protein concentraion (mg/ml) = difference between UV absorbance at 274 nm and at 310 nm, divided by path length of the cell".

xii. 100 µl of the sample were taken in an eppendorf before Factor Xa treatment to run on SDS-PAGE system next day.

xiii. Factor Xa (bought from NEB) was added to the sample (1 unit of Factor Xa per 0.1 mg of the protein) and incubated at 25° C for 24 hours in order to achieve full cleavage of N-terminal His-Tag or MBP-Tag or His-MBP-tag from the protein and also to separate the light chain and heavy chain by cleaving the Factor Xa linker site between the two chains.

xiv. After Factor Xa treatment, 100 µl of the sample were taken to run on an SDS-PAGE system.

xv. 2 samples were prepared, one before and one after the treatment of Factor Xa.

xvi. Each sample contained 25 µl of protein sample, 25 µl of 4x Nu Page loading buffer and was made up to 100 µl with water.

xvii. 2 more samples were prepared in the same way as in step xiv with 25 µl of protein sample, 25 µl of 4x Nu Page loading buffer, 10 µl of 1 M DTT and were made up to 100 µl with water.

xviii. All the four samples were run on an SDS-PAGE system.

xix. A sample gel picture is shown in Fig 2.13

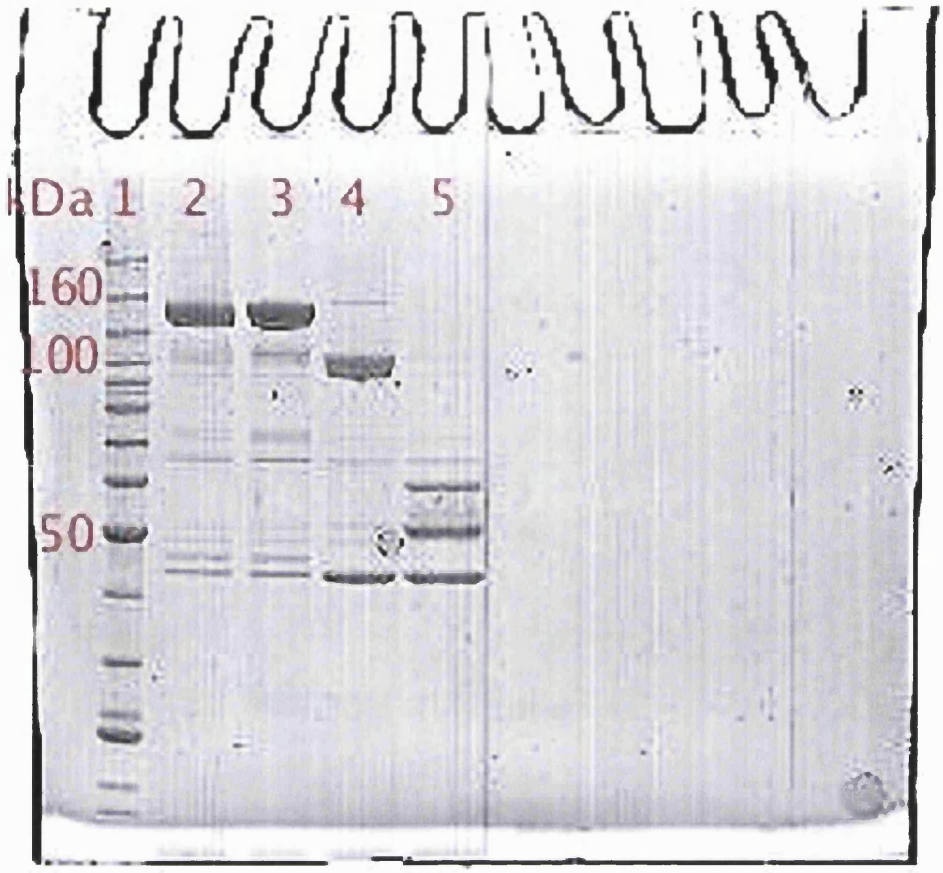


Fig 2.13 SDS-PAGE: LHC-EGF after Factor Xa treatment and reduced with 1M DTT. Lane 1 is the BenchMark™ Protein Ladder (Invitrogen). Lanes 2 and 4: before and after Factor Xa treatment; lanes 3 and 5: before and after Factor Xa treatment plus 1M DTT. The bands observed at a molecular weight of ~150 kDa is the MBP-LHn fusion (lanes 2 and 3). Lane 4 is only LHn after cleaving the N-terminal MBP tag. The cleaved MBP is observed as a strong band at around ~45 kDa molecular weight (lane 4 and 5). In lane 5 the band observed at around ~50 kDa molecular weight is the light and translocation domain of LHn.

Affinity Chromatography 2 :

xx. This step helped in getting more pure protein. After the cleavage of the His-Tag the protein of interest was in the flow through.

xxi. The fractions were run on SDS-PAGE system for verification.

xxii. The buffer used was 0.1 M Tris-HCl pH 8.0, 0.15 M NaCl.

xxiii. A sample gel picture is shown in Fig 2.14.

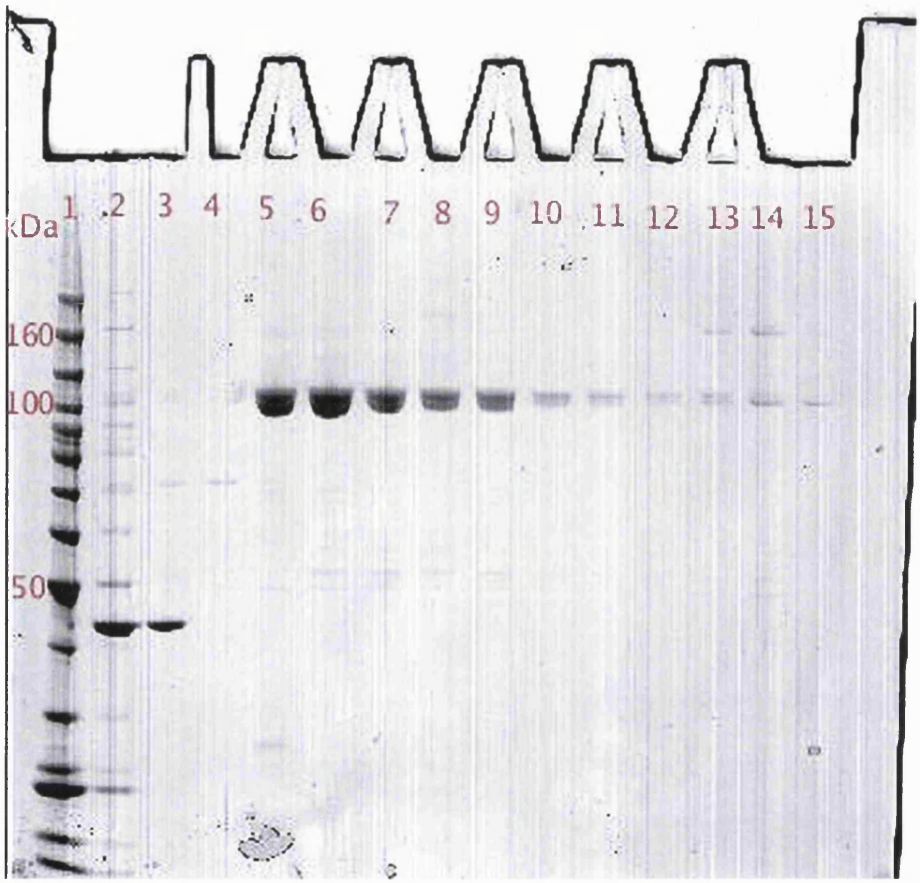


Fig 2.14 SDS-PAGE: LHC-EGF after second affinity chromatography. Lane 1 is the BenchMark™ Protein Ladder (Invitrogen). Lanes 2-15 is flow through. LHC-EGF is seen as a high molecular weight protein (~100 kDa) at the right place on the SDS-PAGE (lanes 5-15).

Size-Exclusion Chromatography:

xxiv. This was the final step in purification of recLHn fragments to get more pure protein.

xxv. Superdex 200 30/100 Pharmacia ready-made prepacked column or Superdex 200 column was used.

xxvi. The buffer used was 0.1 M Tris-HCl pH 8.0, 0.15 M NaCl.

xxvii. After this step the samples were concentrated using 50 kDa molecular weight cut-off concentrators (Vivascience).

xxviii. A sample gel after size-exclusion chromatography is shown in Fig 2.15.

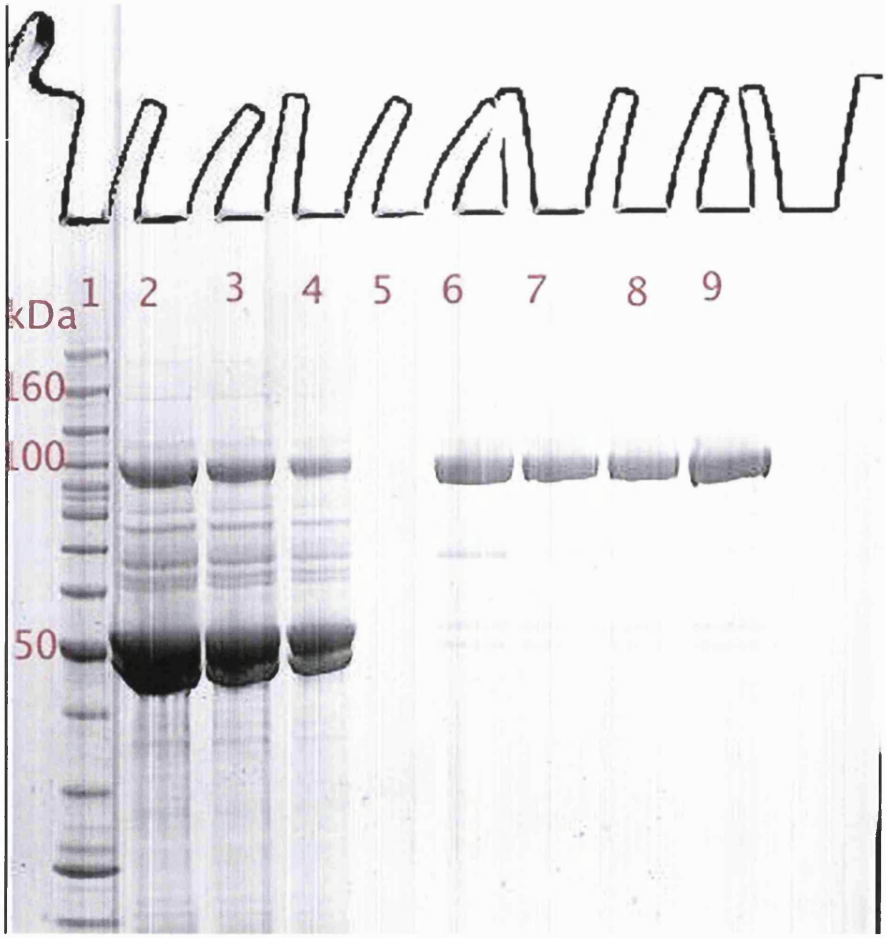


Fig 2.15 A sample SDS-PAGE gel picture after size-exclusion chromatography. Lane 1 is the BenchMark™ Protein Ladder (Invitrogen). Lanes 2-4: concentrated sample after second affinity chromatography; lanes 6-9: after size-exclusion chromatography. The very strong band observed at ~50 kDa is the associated *E.coli* protein that could not be avoided during affinity chromatography. After size-exclusion chromatography the very strong band at 50 kDa (lanes 2-4) is mostly avoided (lanes 6-9). The band at molecular weight ~100 kDa (lanes 6-9) is the LHn.

Final protein concentration and purity assessment:

The protein sample eluted from size-exclusion chromatography was concentrated to desired volume using the 50 kDa molecular weight cutt-off conenctrators. The concentration of the protein sample was measured using UV absorbance (Fig 2.16a) and/or BCA assay (Fig 2.16b).

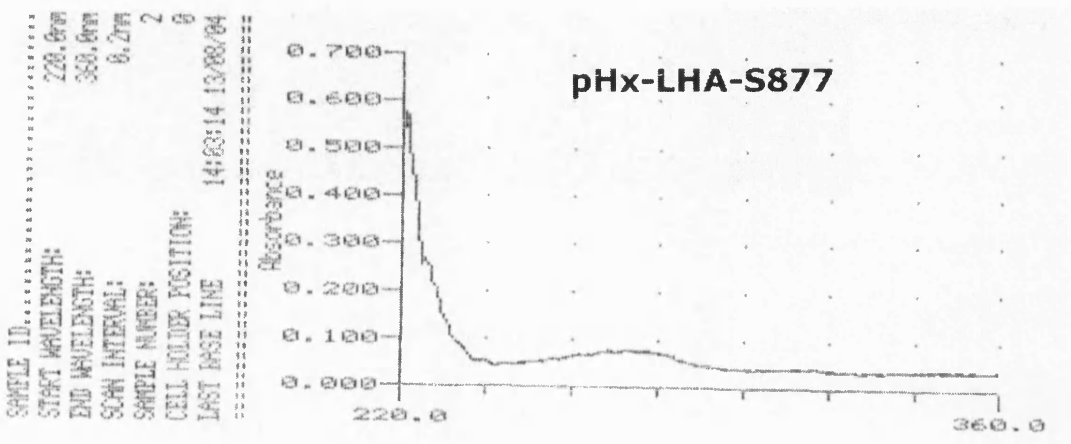


Fig 2.16a UV absorbance spectrum of pHx-LHA-S877.

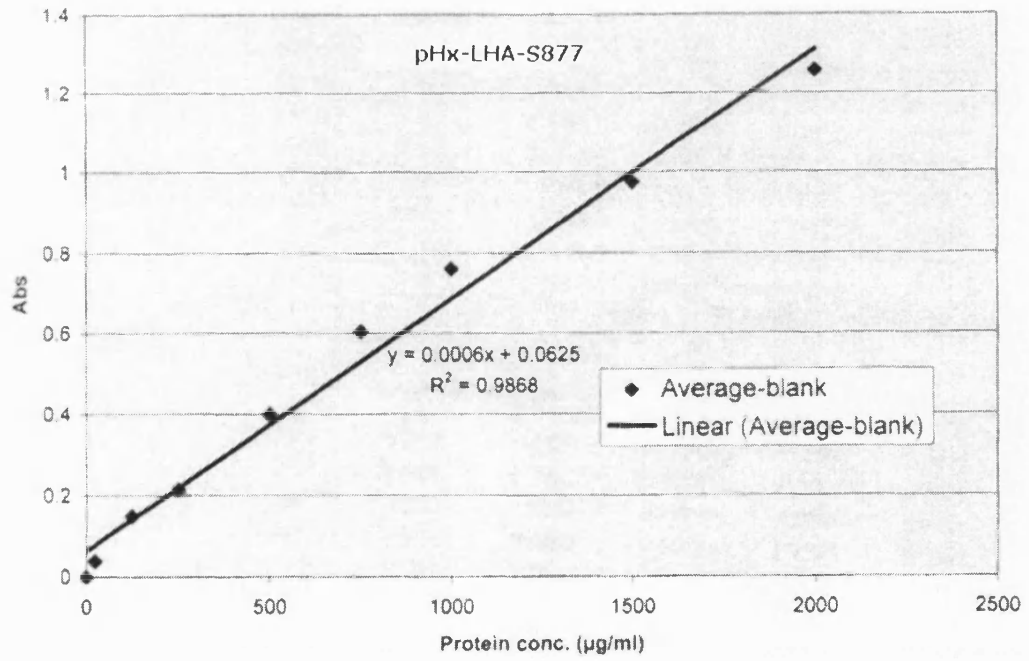


Fig 2.16b BCA assay for pHx-LHA-S877. The mean absorbance for pHx-LHA-S877 sample was 0.875. Substituting this value for *y* in the equation on the graph the concentration of pHx-LHA-S877 was calculated to be 1.35 mg/ml. This was from a litre culture and about ~83 % pure.

For purity assessment the final concentrated sample was loaded on a SDS-PAGE at very high concentration (~0.1 – 0.2 mg/ml). The gel was stained and scanned using the densitometer and a graph was tabulated (Table 2.5) to determine the percentage of purity of the sample (Fig 2.24).

Table 2.3 Densitometer reading for pHMx-LHA-S877 after size-exclusion chromatography and concentration. The sensitivity was increased to 2 instead of 4, which was used as standard for LHn fragments. This increase in sensitivity accounted for the extra bands observed on the gel (Fig 2.17) and is used for other LHn fragments purity assessment.

LINE- 1 Line Width: 3 Pixels

Sensitivity	Noise	Min. Area	Max. Area	Auto Noise	Baseline	Kernel
2	0.001			Yes	Automatic	5
Peak #	Area	Height	Percent	Apex(pixel)	Separation	
1	1.698	0.126	84.487	55	VV	
2	0.029	0.007	1.443	99	BB	
3	0.17	0.018	8.433	131	BB	
4	0.069	0.01	3.423	195	BV	
5	0.045	0.007	2.214	212	BB	

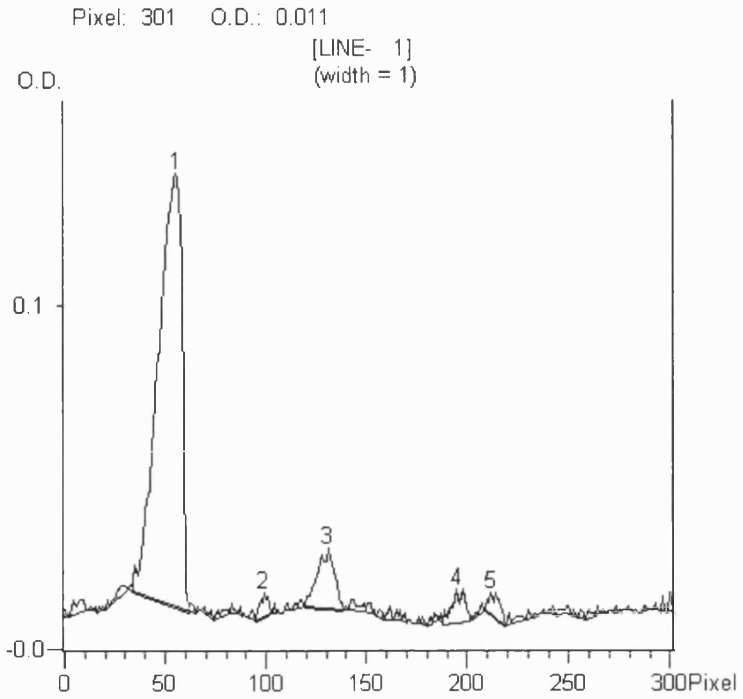


Fig 2.17 Optical density of the stained protein sample – output from the densitometer.

Summary

All the following eleven constructs were cloned, expressed and purification trials were performed.

Table 2.4 List of LHn constructs and protein yeild

LHn plasmid	Molecular weight (~kDa)	Protein yield (mg/l)	Purity (~%)
pHx-LHA-S877	100	0.3 – 3.0	83.0
pHx-LHA-K625A	100	<1.0	<60.0
pHx-LHA-I626A	100	<1.0	<60.0
pHx-LHA-K625A-I626A	100	<1.0	<60.0
pHMx-LHA-S877	100	0.3 – 3.0	94.0

pHMx-LHA-I830 (Truncated LHn/A)	92	Could not estimate the quantity because of poor expression and impurities coordinated during purification process	Not estimated
pMxH-LCC	50	<1.0	<80
pHMx-LHC-His-Tag (pHMx-LHC-HT)	100	35.0 – 40.0	94.0
pMxH-LHC-EGF	107	35.0 – 40.0	94.0
pHMx-LHA-CKC-S877 (V793C/L845K/T865C)	100	0.3 – 3.0	<80
pHMx-LHA-CNC-S877 (V793C/L845N/T865C)	100	0.3 – 3.0	<80

Success has been achieved in obtaining the first X-ray diffraction pattern for LHA-S877 and LHC-EGF. Many different constructs (section 2.8) of LHn fragments were cloned, expressed and purified in order to find a successful candidate for structural studies. Surface engineering of LHn/A, truncation product of LHn/A, light chain of serotype C, and LHC-EGF fusion protein were also expressed and purified. The other constructs were not produced on a large scale and screened for crystallisation because of initial success with LHA-S877 and LHC-EGF. The yield of LHA-S877 is very low ranging between 0.8 mg – 3 mg per litre culture with a purity of ~80%. Hence, LHA-S877 were expressed in large quantities and purified for crystallisation trials with a purity of ~93 %. LHC-EGF is a very well expressed protein about 30 mg – 40 mg per litre culture. Due to time constraint and funding the work is stopped at this stage. Further optimisation of crystallisation conditions LHA-S877, LHC-EGF and LHC-HT is required.

2.9 First Crystal Structure of Recombinant Botulinum Neurotoxin B

Clostridium botulinum B protein receptors in cell-mediated endocytosis are synaptotagmin I and II. The receptors bind to a conserved region in the H_C domain that corresponds to the sialic acid binding site on tetanus toxin. Based on the extensive research performed by our collaborators Dr. Thomas Binz and Dr. Andreas Rummel's laboratory, Institut für Biochemie, Hannover, Germany, attempts were made to co-crystallise/soak the synaptotagmin peptide with botulinum neurotoxin B.

Here we report the first crystal structure of recombinant botulinum neurotoxin B (hereafter referred as BotB) at 1.9 Å resolution (Thiyagarajan et al., unpublished results).

2.9.1 Materials and Methods

Recombinant BotB was prepared (Rummel et al. 2004) and supplied for crystallisation trials at very high concentration of ~10-13 mg/ml. Synaptotagmin II (hereafter referred as Syt2) fragments were mixed to BotB at different ratios to ensure the peptide is bound to BotB. The protein was stored in buffer 50 mM Tris-HCl pH 8.0, 150 mM NaCl, and 2.5 mM $CaCl_2$.

2.9.2 Co-crystallisation of BotB – Synaptotagmin Complex

Attempts were made to co-crystallise BotB-Syt2 complex. Different ratios of Syt2 to BotB were mixed and several hundreds of crystallisation trials were screened. Soaks of Syt2 to the obtained BotB crystals were also tried and several dataset has been collected at Synchrotron Radiation Source at Daresbury Laboratory. Attempts to obtain a BotB-Syt2 complex data were unsuccessful. Recombinant whole length (~1310 aa) BotB was crystallised completely in a different condition those used and reported for crystallisation

of BotBs from native source (Swaminathan and Eswaramoorthy 2000), where ammonium sulphate was used as precipitant. BotB crystallised in $P2_1$ space group using the condition 8-10 % PEG 6000, 0.1-0.25 M Mes buffer at pH 5.5-6.5.

Crystallisation screens were attempted using hanging drop vapour diffusion technique in 24-well Linbrow plates (Molecular Dimensions Ltd, UK). The drop ratio of protein (BotB-Syt2 mixture) to mother liquor is maintained to be 1:1. Different ratios of protein to mother liquor were also tried. Standard structure screens I and II, and PEG/Ion screen from Molecular Dimensions Ltd, Ammonium sulphate grid screen from Hampton Research, were initially screened. The drops were equilibrated with 500-1000 µl of mother liquor in the reservoir and incubated at 16°C. The drops were routinely monitored for any crystallogenesis to appear in the drop after every 24 hours from initial setup.

Observations made directed towards pH range from 5.0 to 7.0 to be screened further using PEG as precipitant. Different PEGs from range 1000 to 12000 were tried. Buffers tried were Bis-Tris, Mes, Sodium cacodylate, and Na citrate. All chemicals were purchased from Sigma (UK), PEGs from Fluka (UK). High-quality grade HPLC water is used to prepare all the solutions. Crystalloids were observed after a week in drops containing PEG 6000 in buffer Mes pH 6.0. Varying the precipitant concentration against constant ionic strength resulted in thin plate crystals, needles, and crystalloids and of not use for X-ray diffraction studies. Attempts made in varying the concentration of the buffer resulted in better crystals. These crystals diffracted to 2.6 Å resolution at the SRS, Daresbury.

Optimisation was carried to grow better crystals in order to obtain higher resolution data. Several additives like nickel chloride, N-acetyl galactosamine, N-acetyl glucosamine, alcohols such as MPD (2-methyl-2,4-pentanediol), ethylene glycol, glucose, sucrose were tried. Detergents were also used as additives in crystallisation. BotB-Syt2 were mixed at equal proportions and

screened for crystallisation using the same condition 8-10 % PEG 6000, 0.1-0.25 M Mes pH 5.5-6.5 whilst detergents were used as additives this time. Parallelopiped single crystals grew after a week in a drop containing 1:1:0.3 ratio of protein to mother liquor to detergent equilibrated against the reservoir solution 8-10 % PEG 6000, 0.1-0.25 M Mes pH 5.5-6.5. A single crystal grown using sucrose monolaurate detergent as additive diffracted to 1.9 Å resolution and data were collected at PX station 14.2, SRS, Daresbury. The toxin crystals lose their diffraction quality though not the shape after three weeks from they are formed. Hence, the crystals need to be freshly grown prior to data collection.

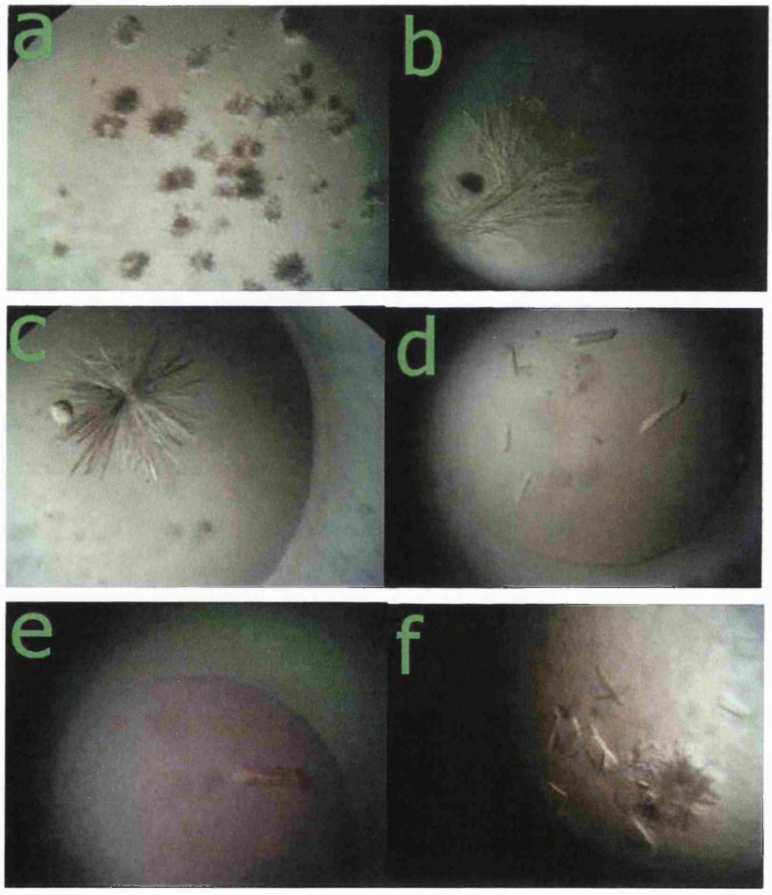


Fig 2.18 Different crystals of BotB: a) 12 % PEG 6000, 0.1 M Mes pH 6.0; b) 10 % PEG 6000, 0.1 M Mes pH 6.0; c) 8% PEG 6000, 0.15 M Mes pH 6.0; d) 10 % PEG 6000, 0.25 M Mes pH 6.0; e) 10 % PEG 6000, 0.25 M Mes pH 6.0, sucrose monolaurate; f) 10% PEG 6000, 0.25 M Mes pH 6.0, 2mM NAG.

2.9.3 Single Crystal Data Collection and Processing

X-ray diffraction data were collected at ~100K at SRS, Daresbury Laboratory on PX station 14.2 (λ=0.978 Å). Diffraction images were recorded on an ADSC Quantum 4 Detector. The crystal was cryoprotected in 25 % Glycerol in mother liquor prior to data collection. Diffraction data were collected with an oscillation range of 1°.

HKL2000 (Otwinowski and Minor 1997a) suite was used for X-ray data processing and scaling. The data were indexed, processed and scaled in primitive monoclinic space group with cell dimensions a=76.24 Å, b= 122.79 Å, c= 95.19 Å, α= 90°, β=112.4°, γ=90°. Calculation of Matthew's coefficient gave a value of V_m=2.96 with 1 molecule per asymmetric unit. Initial indexing resulted in a space ambiguity between C-centred monoclinic and primitive monoclinic. There was doubling of two cell dimensions in C-centred monoclinic space group compared to primitive monoclinic. This could be because of the presence of pseudocentering in C-centered space group. Initially data were processed in C-centred monoclinic space group because of excitement that the BotB-Syt2 existence might have given a new crystal form. But after molecular replacement solution, and electron density calculation, and analysis using graphics disproved the choice of selection of C-centred monoclinic system. The data was reprocessed in primitive monoclinic space group P 2_1.

2.10 Results and discussion

Molecular Replacement

Initial phases were obtained using molecular replacement method from the phases of 1EPW with protein atoms only. The program AmoRe or Molrep (Navaza 1994; Vagin and Teplyakov 1997) was used for molecular replacement method. Phases were refined using Crystallographic and NMR

suite (CNS) with a test set of reflections kept aside to calculate the R_{free} (Brunger 1992; Brünger et al. 1998). Crystallographic data collection statistics are given in Table 2.5.

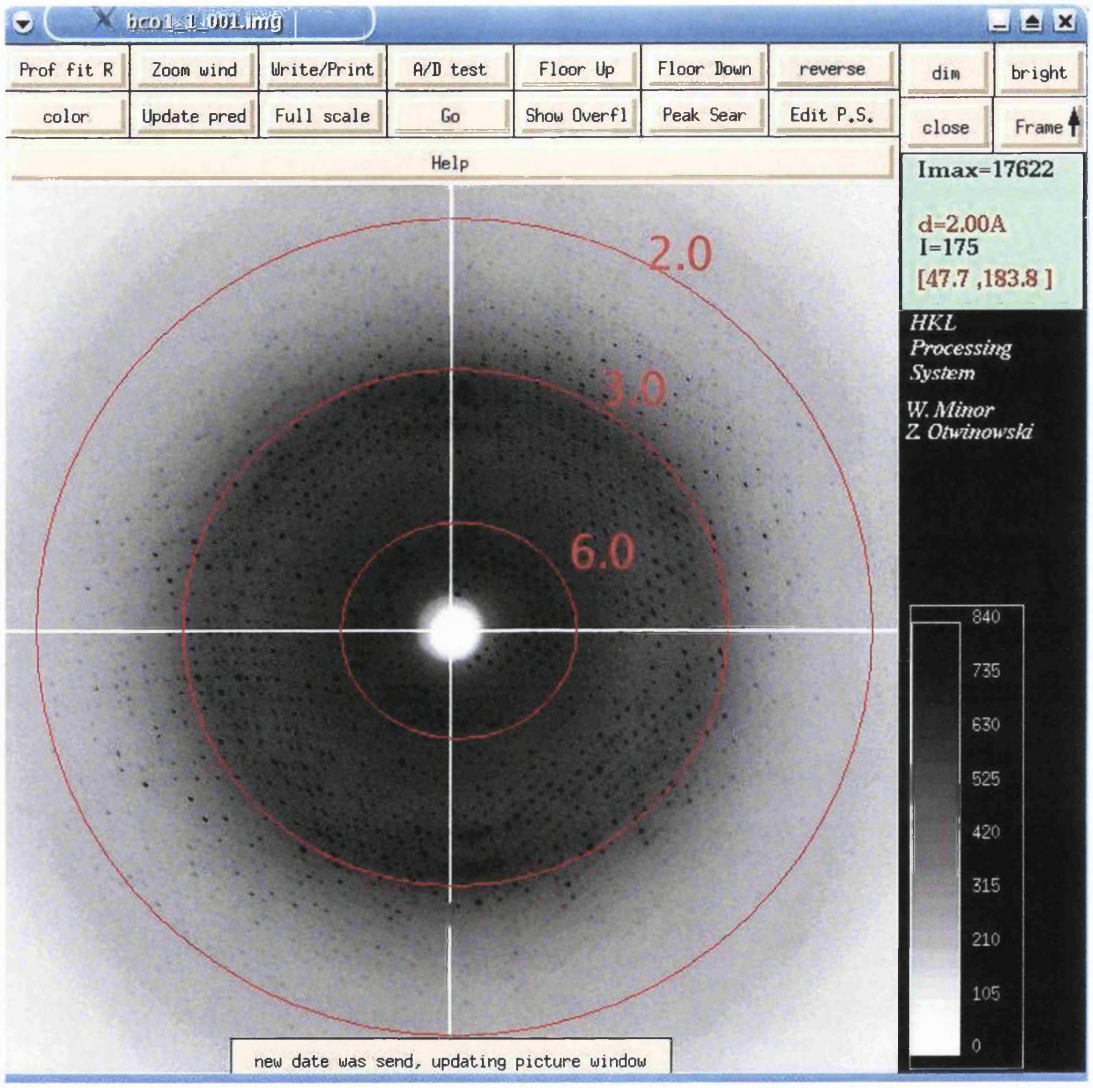


Fig 2.19 X-ray diffraction image of BotB. The resolutions circles are shown in red. The numbers on the resolution circles refers to the resolution of X-rays in angstrom units.

Table 2.5 Crystallographic data collection and refinement statistics

	BotB
Resolution (Å)	50-1.9
(Outermost shell) (Å)	1.97 – 1.90
Reflections measured	1149513
Unique reflections	129075
R_{symm}[a]	0.12
(Outermost shell)	0.70
Completeness	90.5
(outermost shell) (%)	52.3
<I/σI> (outermost shell)	8.77 (1.14)
R_{cryst}[b]	0.25
R_{free}[c]	0.31
Number of solvent molecules	882
RMS deviation from ideality	
in bond lengths (Å)	0.008
in angles (°)	1.4
Average B factor (Å^2)	31.8
Protein Atoms	31.7
Solvent	31.4
Zn^{2+}	22.8
Ca^{2+}	34.8

[a]$R_{symm} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I\rangle(hkl)|/\Sigma_{hkl}\Sigma_i I_i(hkl)$ where <I> is the averaged intensity of the *i* observations of reflection *hkl*.

[b]$R_{cryst} = \Sigma||F_O| - |F_C||/\Sigma|F_O|$ where F_O and F_C are the observed and calculated structure factor amplitudes, respectively.

[c]R_{free} is equal to R_{cryst} for a randomly selected 5% subset of reflections not used in the refinement.

Structure Analysis and Validation

For the first time the recombinant BotB was crystallised with the full length of the toxin with the linker region remain intact (Fig 2.20) and not proteolytically cleaved between light and heavy chain domains. The disulphide bridge between residues Cys 441 and Cys 450 is clearly mapped in the electron density. The N-terminal portion of the molecule was less disordered but most of the backbone and sidechain positions could be fitted into the electron density map. Phases for the C-terminal residues from 1290 are highly disordered and could not be modelled into the electron density map.

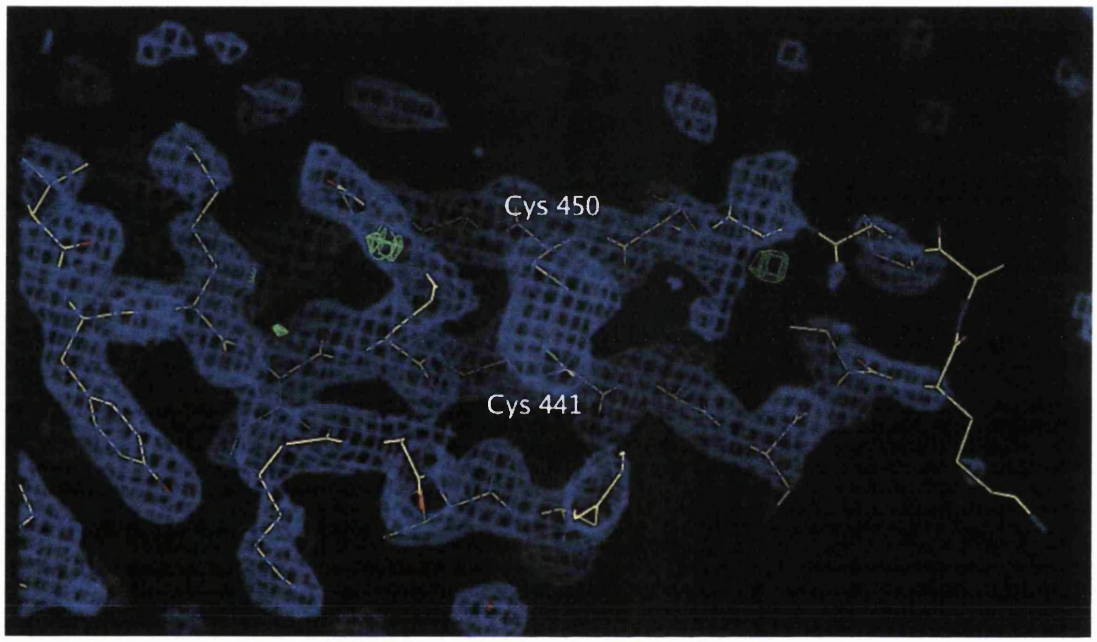


Fig 2.20 $2F_O$-F_C (in blue contoured at 1.0 σ) and F_O-F_C (in green contoured at 3.0 σ) electron density maps for the linker region is shown. The average thermal factor for atoms in this loop is around 55.

After the first round of refinement Zn^{2+} ion was inserted at the active site with the help of $2F_O$-F_C and F_O-F_C weighted SIGMAA maps in its coordination sphere (Fig 2.21). The translocation belt partially occludes the Zn^{2+} binding site similar to 1EPW (Swaminathan and Eswaramoorthy 2000). After careful analysis of the whole toxin structure the presence of Ca^{2+} ion is observed near

the hinge between translocation domain and the catalytic domain (Fig 2.23. Though crystallisation conditions did not include any Ca^{2+}, it might have been carried through during the purification process of the protein. Two glycerol molecules and one sulphate ion were also found on the surface of the toxin. Glycerol molecules are found one at the C-terminal domain interacting with main chain atoms of residues G1108, N1109 and I1244 and one near the light chain interacting with ND2 of N134, whereas the sulphate molecule was found near the hinge between light chain and translocation domain.

Role of zinc ion:

Extensive research has been done and reported on the role of Zn^{2+} ion in botulinum neurotoxins (Vallee and Auld 1990a; b; Montecucco and Schiavo 1993). Zn^{2+} ion is very important for the catalytic activity of the toxin, though substitution with other metal ions (Eswaramoorthy et al. 2004) could render some catalytic activity to the enzyme. It is thought that the presence of Zn^{2+} ion could be for both structural and functional stability of the molecule, but reports have shown that the presence of zinc ion is only catalytic and has no influence on the conformation of the light chain (Breidenbach and Brunger 2004; Eswaramoorthy et al. 2004; Agarwal et al. 2005).

Role of calcium ion in botulinum neurotoxins:

Presence of Ca^{2+} ion in BotB participates in intracellular translocation of the toxin (Eswaramoorthy et al. 2004). The calcium ion coordinates with three light chain residues P281, I284 and D289, and one residue N488 from the translocation domain belt region (Table 2.6) (Fig 2.22). There is also a coordinating water molecule present at the calcium bound site. Presence of calcium ion in any other serotype of botulinum neurotoxins is not reported yet. The presence of calcium ion is common among zinc proteases like thermolysin (Colman et al. 1972; Matthews et al. 1974), adamalysin II (Gomis-

Ruth et al. 1994), and matrix metalloproteases. The presence of calcium ions is thought to play a role in structural stability of the enzyme. About 20-30 % of botulinum neurotoxins light chain is restored by metal substituent like Mg, Mn or Ca (Ahmed and Smith 2000).

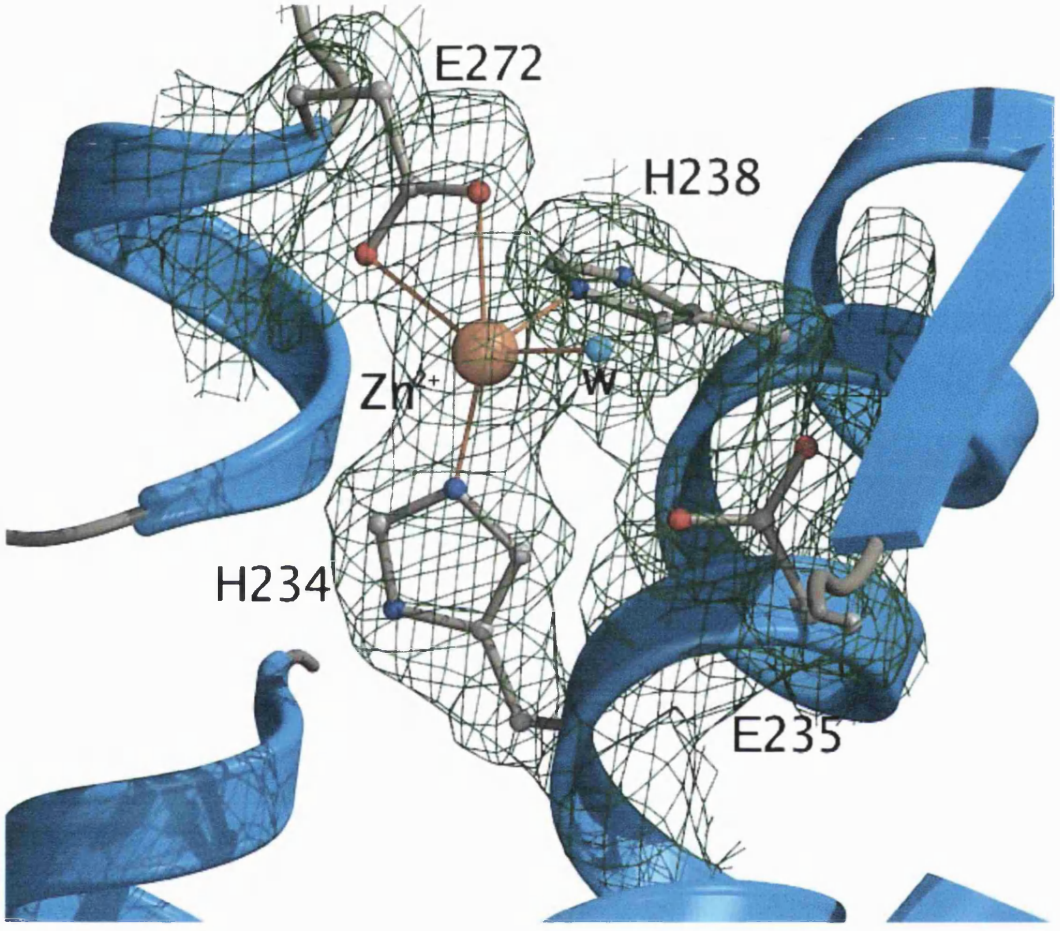


Fig 2.21 Zn^{2+} site at the catalytic centre. Zinc ion – *gold*; water – *cyan*; coordinating residues in ball-and-stick model – *grey* and standard atom colours for atoms belonging to the protein residues. $2F_O$-F_C electron density map contoured at 1.0 σ.

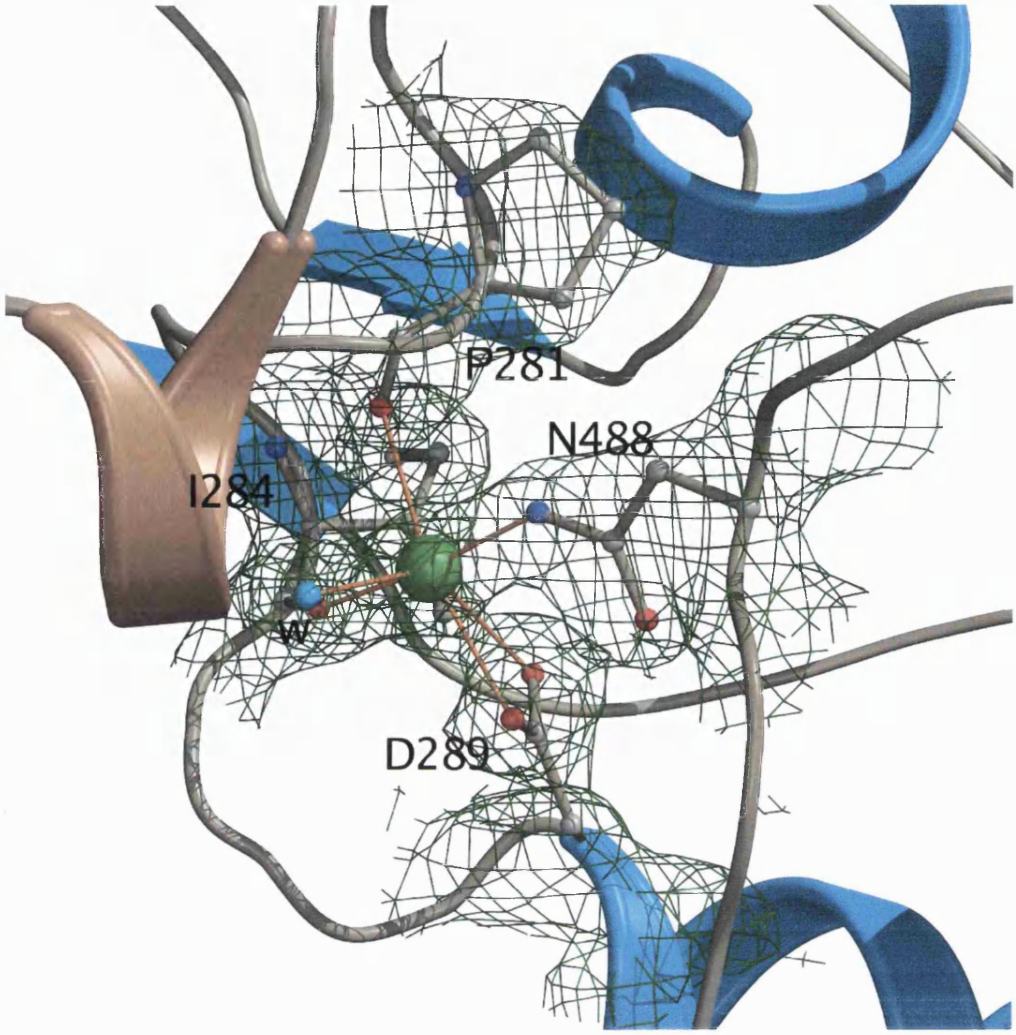


Fig 2.22 Ca^{2+} ion between translocation domain and catalytic domain. Calcium ion – *green*; water – *cyan*; coordinating residues in ball-and-stick model – *grey* and standard atom colours for atoms belonging to the protein residues. $2F_O$-F_C electron density map contoured at 1.0 σ.

Analysis of BotB structure with available botulinum B structure (PDB: 1EPW) (Swaminathan and Eswaramoorthy 2000) superpose with a root-mean square deviation (rmsd) value of 0.62 Å for C-alpha atoms. The catalytic domain (LC) (light chain) comprises of a mixture of α-helices and β-strand secondary structures with the coordinating metal ligand Zn^{2+} ion at the active site (Fig 2.21). The active site is located in a deep cleft partially occluded by the translocation belt of the translocation domain. The Zn^{2+} coordination and conserved residues around this site is well pronounced in all CNTs.

Table 2.6 Coordination distance for Zn^{2+} and Ca^{2+} ions in BotB

Metal Ion	Amino Acid Side Chain	Distance (Å)
Zn^{2+}	His 234 NE2	2.10
	His 238 NE2	2.23
	water	2.42
	Glu 272 OE1	2.33
	Glu 272 OE2	2.42
Ca^{2+}	Pro 281 O	2.56
	Ile 284 O	2.44
	water	2.45
	Asp 289 OD1	2.50
	Asp 289 OD2	2.56
	Asn 488 ND2	2.40

The translocation domain (H_N) has its unique structural feature with two long amphipathic α-helices that run anti-parallel to each other (Fig 2.23). Cys 450 from translocation domain forms an intramolecular disulphide bridge with Cys 441 of light chain. The translocation belt from E487 to F542 spans the catalytic domain partially occluding the zinc site inaccessible to the substrate and thereby rendering the toxin inactive. The translocation domain is believed to form ion-conducting channels in artificial membranes (Donovan and Middlebrook 1986; Blaustein et al. 1987; Shone et al. 1987; Schmid et al. 1993). BotB reported here is from a recombinant source and is not proteolytically cleaved between light chain and heavy chain and the linker region (residues 440-443) is intact.

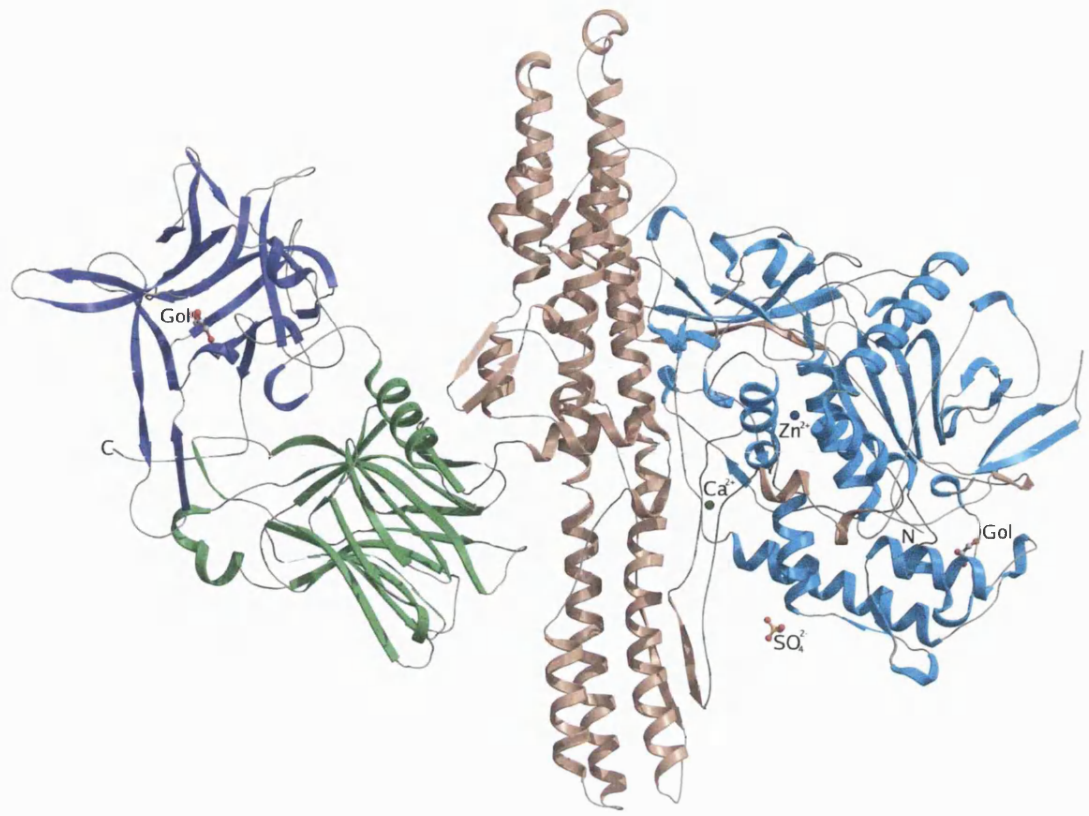


Fig 2.23 Ribbon diagram representation of whole toxin recombinant BotB. Light chain (LC) – cyan; translocation domain H_N – brown; Binding domain – H_C: H_{CN} – green; H_{CC} – violet; Zinc – blue; Calcium – green; sulphate and glycerol molecules are shown as ball-and-stick models

The binding domain (H_C) has two sub-domains mainly of β-sheet topology linked by a single α-helix. The C-terminal residues play a very important role in receptor binding. Mutation studies of the C-terminal residues resulted in loss of activity of the toxin or intoxification into the neuronal cells resulting in failure of binding of the toxin to the cell membrane. The superposition of the binding domain of BotB with 1F31 in complex with sialyllactose reveals glycerol molecule of BotB bound near the ganglioside binding region of the binding domain (Fig 2.24) (Swaminathan and Eswaramoorthy 2000).

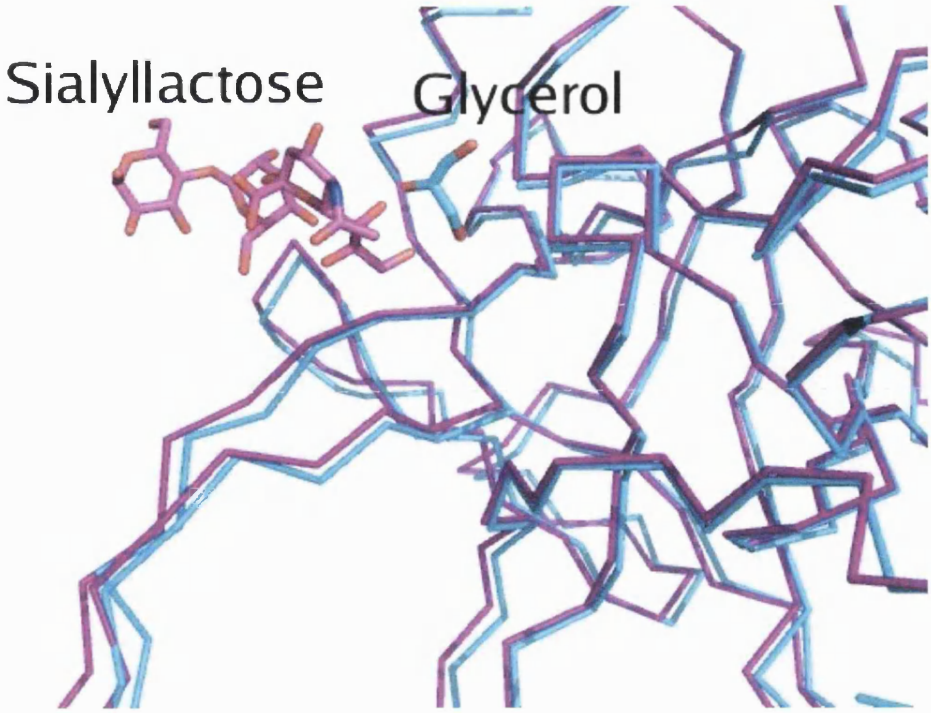


Fig 2.24 Superposition of binding domains of BotB – *cyan* and 1F31 – *magenta* (Swaminathan and Eswaramoorthy 2000). Corresponding ligands are shown in similar colour with respect to the molecule they belong to.

Analysis of structural data confirms that more than 85.5 % of residues in most favourable region, 12.6 % in additionally allowed region, and 1.3 % in generously allowed region of the Ramachandran plot (Fig 2.25), compared to only 0.6 % of residues in disallowed region. This is because of poor or disordered electron density map where the regions could not be fixed.

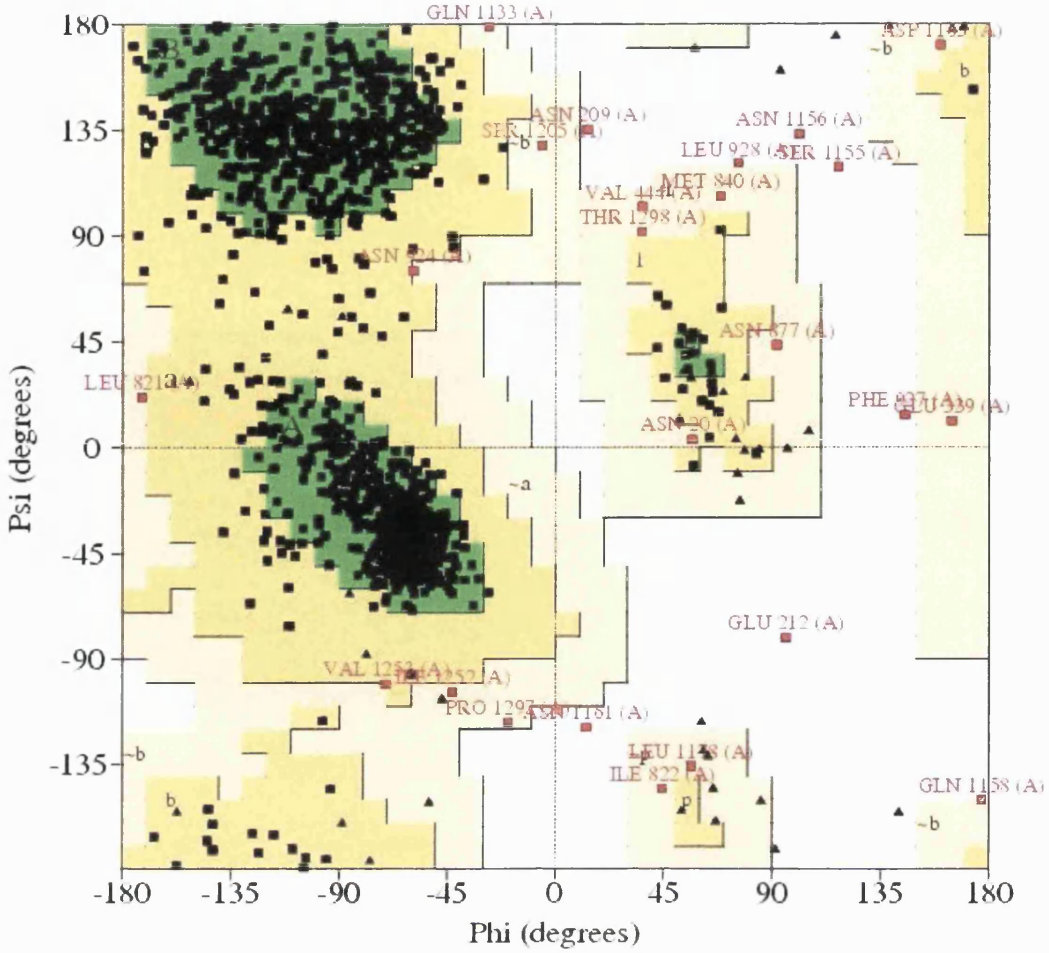


Fig 2.25 Main chain torsion angle plot (Ramachandran plot) for BotB

Conclusion

Several attempts to co-crystallise BotB-Syt2 complex in equi- and/or variable proportional mixtures were unsuccessful. This could be because of crystallographic artifact where the toxin molecules pack in a particular fashion. Thus the complexes in solution might have fallen apart under crystallographic conditions. A glycerol molecule was found at the substrate-binding site of botulinum neurotoxin B (Fig 2.24). It is evident that if synaptotagmin II were bound at the ganglioside binding site of BotB there would not be any/enough space left for the glycerol molecule to be seen at the site. This explains the fact that, though the BotB-Syt2 is found as complex in solution but not in crystals.

One way to overcome this problem is to express and purify the protein and peptide as a fusion protein using recombinant methods. The peptide could be linked to the C-terminal of the protein by a short flexible linker molecule/peptide of length ~10 - 20 amino acids, thus allowing the ~65 amino acid length synaptotagmin II to span freely over the ganglioside binding domain. Other approach could be to purify and crystallise only the binding domain of BotB with synaptotagmin II. This might help the complex to be crystallised without any fusion approach.

Chapter 3

Crystal Structures of *Erythrina cristagalli* Lectin with Bound *N*-linked Oligosaccharide and Lactose

Introduction

Lectins are carbohydrate-binding proteins (glycoprotein) of molecular weight ranging between 60 kDa - 110 kDa and are often classified based on saccharide specificity. Lectins can be found in plants, animals, bacteria and viruses and are involved in biological recognition functions. Lectins have one or more binding sites per subunit, which can reversibly bind to specific sugar segments through hydrogen bonds and van der Waals interactions (Ashford et al. 1991). The term lectin (legere - Latin verb for to select) was coined by Boyd to emphasise the ability of some hemagglutinins to discriminate blood cells within the ABO blood group system (Boyd and Reguera 1949). Some of the lectins were found to be toxic to humans (Vasconcelos and Oliveira 2004) and animals. In general, nausea, bloating, vomiting and diarrhea characterize the oral acute toxicity of lectins on humans exposed to them. Plant lectins are subdivided into four major classes: Merolectins, Hololectins, Chimerolectins and Superlectins (Van Damme et al. 1998).

Depending on the specificity towards a given monosaccharide and/or oligosaccharide the lectins will selectively bind to one of these sugars (D-mannose/D-glucose, *N*-acetyl-D-galactosamine, *N*-acetyl-D-glucosamine, L-fucose and N-acetylneuraminic acid) that are typical constituents of eukaryotic cell surfaces (Ashford et al. 1987). One of the most important features of plant lectins is their defense mechanism against proteolysis and stable over a wide pH range. Legume lectins serve as a model system to study the molecular basis of protein-carbohydrate specificity. The specificity, high stability over a wide pH range, etc. has attracted lectins towards therapeutics as a fusion protein with botulinum neurotoxin in the treatment of chronic pain(Chaddock et al. 2004).

3.1 *Erythrina cristagalli* lectin

Fig 3.1 *Erythrina cristagalli*

Erythrina cristagalli is an ornamental plant commonly called as 'coral tree' (Fig 3.1). The lectin from this plant is specific for galactose and *N*-acetyl-galactosamine (Iglesias et al. 1982; Ashford et al. 1991). *Erythrina cristagalli* lectin (ECL) exhibits similar affinity to fucosyllactose and fucosyllactosamine, whereas other members show higher affinity for fucosyllactose (Teneberg et al. 1994; Moreno et al. 1997; Srinivas et al. 2001; Svensson et al. 2002). *in vitro* studies on ECL has shown that they have hemagglutinating activity and is mitogenic for human T lymphocytes (Iglesias et al. 1982; De Boeck et al. 1985; Harris et al. 1987). Recently, a conjugate comprising a catalytically active derivative of *Clostridium botulinum* neurotoxin A (LH_N/A) coupled to ECL has been used to selectively target nociceptive afferents (Duggan et al. 2002; Chaddock et al. 2004).

3.2 Medical application of ECL

ECL due to its high carbohydrate specificity has drawn the attention of pharmaceutics and researchers to use this as a drug delivery vehicle for use in gene therapy to cure SCID (Severe Combined Immunodeficiency Disease) and in the treatment of chronic pain (Chaddock et al. 2004).

3.3 Materials and methods

Crystallisation, data collection and structure solution of ECL were performed by Dr. Kathryn Turton. Model building, structure validation and analysis were performed by me.

Protein purification and crystallisation: Native ECL (*n*ECL) was purchased from Sigma (UK). The purified recombinant ECL (*rec*ECL) was supplied from Dr. John Chaddock's laboratory. The DNA extracted from *E.cristagalli* leaves were amplified by polymerase chain reaction (PCR) and cloned into vector pMTL1015. The clone was transformed into *E.coli* BL21 (DE3) cells and expressed (OD_{600} = 30) in an 8-litre fermentation culture at 37° C. The protein was solubilized from inclusion bodies, subsequently refolded and purified on an immobilized lactose matrix. After extensive washing, *rec*ECL was eluted by the addition of 0.3M lactose (Stancombe et al. 2003).

Crystallisation was achieved using the vapour diffusion method (hanging drops) at 16°C and crystals were observed within 3-4 weeks. The crystallisation conditions used is given in Table 3.1.

Table 3.1 Crystallisation conditions used to crystallize *n*ECL and *rec*ECL.

Protein	Crystallisation condition	
	Drop (4µl)	**Reservoir**
*n*ECL (space group $P6_5$)	2µl of protein solution, 2µl of mother liquor plus 0.4µl of 100 mM lactose solution.	800µl of mother liquor (70% MPD and 0.1M HEPES, pH 7.0).
*rec*ECL (space group P1)	2µl protein and 2µl of mother liquor.	800µl of mother liquor (17% PEG 3350, 0.3M sodium chloride and 0.02M imidazole).
*rec*ECL (space group $P2_1$)	2µl protein and 2µl of mother liquor.	800µl of mother liquor (20% PEG 3350 and 0.2M imidazole).

X-ray data collection and structure determination: X-ray diffraction data were collected at 100K using the Synchrotron radiation source at Daresbury, UK on station PX14.1 (wavelength 1.483Å). Raw data images for *n*ECL were indexed and integrated using DENZO (Otwinowski and Minor 1997a), then scaled with SCALEPACK (Otwinowski and Minor 1997a). X-ray images collected for the *rec*ECL crystals were processed and scaled using HKL2000 (Otwinowski and Minor 1997a) (Table 3.1).

Phase determination: Initial phases were obtained by the molecular replacement method using the program AMoRe (Navaza 1994). For *n*ECL, the crystal structure of ECorL at 1.95Å resolution (PDB code 1AX1) (Elgavish and Shaanan 1998) was used as an initial model with the nine non-conserved residues mutated to alanine and all waters, ligands and metal ions removed. Refinement was performed using the CNS suite of programs (Brünger et al. 1998). After one round of refinement, the crystallographic R factor (R_{cryst}) dropped to 29.59% (R_{free} 32.38%, based on 5% of reflections omitted from the refinement). Calculated phases from the refined structure were used to determine $|F_o|-|F_c|$ and $2|F_o|-|F_c|$ electron density maps. Careful examination

of the $|F_o|-|F_c|$ map allowed mutation of the non-conserved residues back to their native side-chains and the addition of calcium and manganese ions to the model. HEPES and lactose were built into electron density of each protomer, with parameter and topology files from the HICUP server (Kleywegt and Jones 1998a). Similarly glycosylated sugars (3 in molecule A and 6 in molecule B, part of the heptasaccharide bound to Asn113) were modelled into the $|F_o|-|F_c|$ electron density map, riding on Asn113 residue. Repeated rounds of refinement and model building were carried out to improve the model. After several rounds of refinement, water molecules were added to the structure if there were peaks in the $|F_o|-|F_c|$ electron density maps with heights greater than 3σ at hydrogen bond forming distances from the appropriate atoms. $2|F_o|-|F_c|$ maps were also used to check the consistency in peaks. Water molecules with a temperature factor of >60Å^2 were excluded from subsequent refinement steps.

Subsequent solution of *rec*ECL structures was achieved using the refined coordinates of *n*ECL (at 2.0Å resolution) as a search model. Model building and refinement procedures were undertaken as described for *n*ECL. For *rec*ECL in space group P1, apart from the lactose molecule glycerol was modelled into the carbohydrate-combining site. The program PROCHECK (Laskowski et al. 1993) was used to assess the quality of each structure after the final round of refinement. Analysis of the Ramachandran plot for each structure revealed that over 99% of residues were located in allowed regions of the plot. All the figures were prepared with the program Bobscript (Esnouf 1997a).

3.4 ECL structure

ECL is active in its dimeric state and adopts the conserved jelly-roll topology (Fig 3.2), characteristics of other legume lectins. It forms non-canonical dimers via the handshake motif which was first observed in *Erythrina corallodendron* lectin (EcorL) (Shaanan et al. 1991), a close homolog (96%) to

ECL. In the EcorL structure the *N*-linked heptasaccharide is bound to Asn17, which is at the dimer interface and forms the handshake motif. The carbohydrate moiety from the two molecules in the asymmetric unit interacts by means of hydrogen bonds and is believed to be responsible for the non-canonical dimer formation. But, in the native structure of ECL, the protomers are *N*-glycosylated at Asn113 and the carbohydrate moiety is on the opposite side of the dimer interface (Fig 3.2).

The ECL structure comprises a six-stranded back β-sheet, a short five-membered β-sheet, and a set of loops connecting the three sheets. Each protomer contains one Ca^{2+} and Mn^{2+} ion (Fig 3.3), both of which are required for carbohydrate binding activity (Emmerich et al. 1994). These ions help in maintaining the correct spatial orientation of the carbohydrate binding residues (Derewenda et al. 1989). The interaction of these ions within the molecule in the asymmetric unit is listed in Table 3.2. A *cis*-peptide between residues Ala88 and Asn89 holds the side chain of Asn89 in the correct orientation for carbohydrate binding (Svensson et al. 2002).

Table 3.2 – Contacts between ECL and metal ions

Metal Ion	Amino Acid Side Chain	Distance (Å)
Ca^{2+}	Asp136 OD2	2.40
	Asp129 OD2	2.53
	Asp 129 OD1	2.60
	Asn133 OD1	2.45
	water	2.54
	water	2.55
	Asp129 CG	2.9
	Asp129 OD1	2.6
	Asp129 OD2	2.5
	Phe131 O	2.5
	Asn133 OD1	2.5
Mn^{2+}	His142 NE2	2.39
	Glu127 OE2	2.22
	Asp129 OD2	2.23
	Asp136 OD1	2.18
	water	2.33
	water	3.96
	water	2.21
	Glu127 OE2	2.2
	Asp129 CG	3.2
	Asp129 OD2	2.2
	Asp136 CG	3.1
	Asp136 OD1	2.2
	His142 NE2	2.4
	His142 CE1	3.2

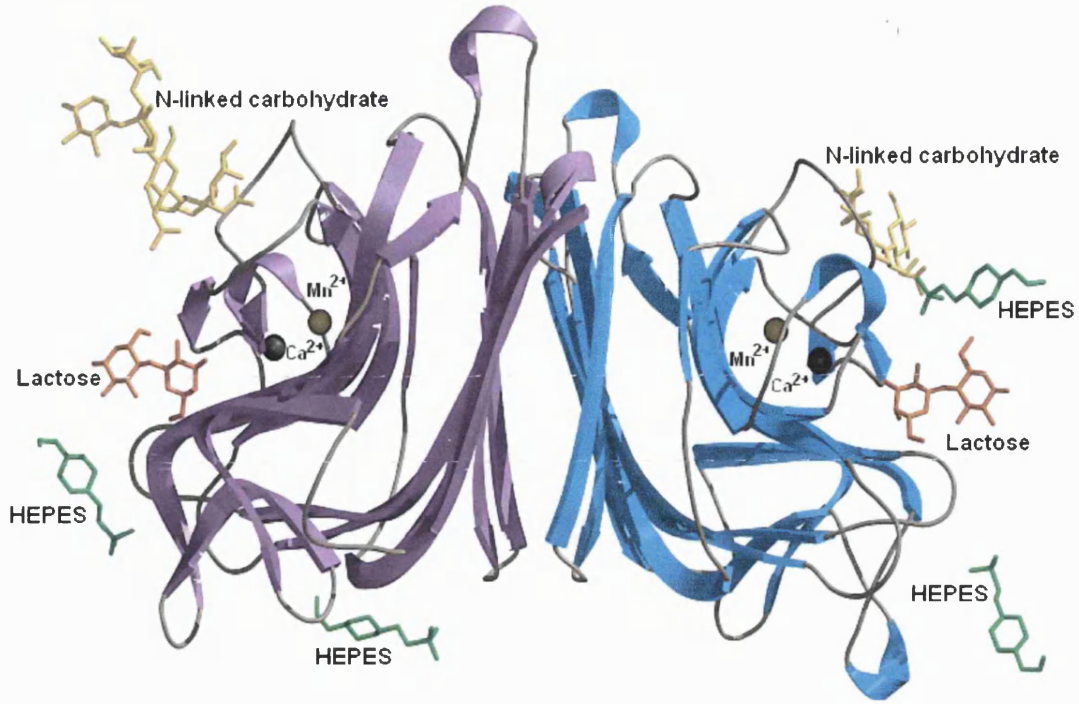


Fig 3.2 The structure of ECL dimer. Protomers of ECL associate together back-to-back to form non-canonical dimers via the "handshake motif" as was first observed for ECorL. The dimer structure is stabilized by two hydrogen bonds and a series of contacts between side chains in four strands of the flat, six-stranded β-sheet. The protomers are tilted with respect to one another such that the N- and C-termini play no part in the dimer interface. Each protomer adopts the conserved jelly-roll fold characteristic of legume lectins. The *N*-linked carbohydrate (yellow) (on Asn113), lactose (gold) bound at the combining site and HEPES (green) molecules (from the crystallisation medium) are shown. The manganese (dark green) and calcium (grey) ions bound in the vicinity of the combing site are shown as small spheres.

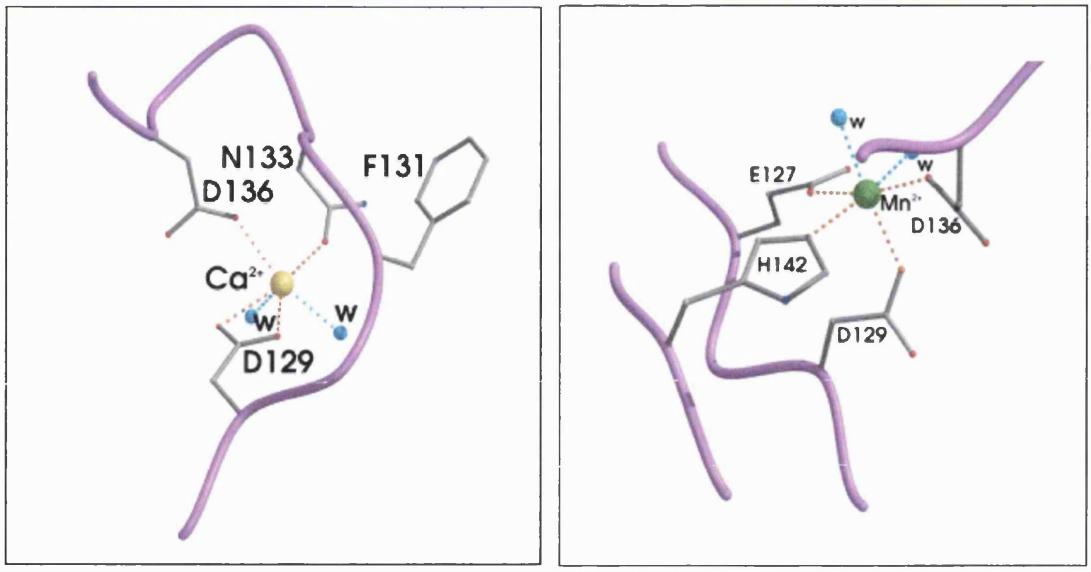


Fig 3.3 Ca^{2+} (yellow) and Mn^{2+} (green) ions interaction with the molecule is shown. Water molecules are shown in blue colour and labelled as 'w'.

3.5 Results and discussion

Native ECL (*n*ECL) - 1UZY: *n*ECL co-crystallized with lactose in space group $P6_5$, with two molecules per crystallographic asymmetric unit and 67% of the crystal volume occupied by solvent. The final model (dimer) at 2.0Å resolution (R_{cryst} 18.97%, R_{free} 20.92%) contains 477 amino acids, two calcium and manganese ions, two lactose molecules, four GlcNAc residues, two fucose residues, one xylose, two mannose residues, four HEPES molecules and 456 water molecules (Table 3.3). The estimated Luzzati coordinate error is 0.22Å. Analysis of the Ramachandran plot revealed that the side-chains of over 99% of the amino acids have allowed or additional allowed conformations. The side-chains of Tyr106 in each molecule adopt generously allowed conformations. The first three residues of the heptasaccharide bound to Asn113 were modelled into the electron density in one of the two protein molecules in the asymmetric unit, while electron density for six of the seven sugar residues was observed in the other.

Recombinant ECL (*rec*ECL): *rec*ECL (1UZZ) crystallized in triclinic form, with four molecules per unit cell (49% of the crystal volume occupied by the solvent). The final model at 2.13Å resolution (R_{cryst} 19.1%, R_{free} 24.1%) contains 958 amino acids (molecule A, 1-239; molecule B, 1-239; molecule C, 1-240 and molecule D, 1-240), four calcium and manganese ions, three glycerol molecules and 879 water molecules (Table 3.3). The Luzzati coordinate error is 0.23Å and the average B factor for protein atoms is 24.52Å^2. $|F_o|-|F_c|$ electron density in the combining site of three protomers was identified as glycerol, which had been included in the cryoprotectant solution used during data collection. A single residue, Tyr106, in only one of the four protomers is located in a generously allowed region of the Ramachandran plot. Possible alternative conformations were observed in the electron density map for the side-chains of Ser120, but not in all of the molecules in the unit cell.

*rec*ECL (1V00) was also co-crystallized with lactose in space group $P2_1$, with four molecules per asymmetric unit (49% solvent). The final model at 1.7Å resolution (R_{cryst} 17.79%, R_{free} 20.31%) contains 959 amino acids, four lactose molecules, four calcium and manganese ions and 1119 water molecules (Table 3.3). The Luzzati coordinate error is 0.20Å and the average B factor for protein atoms is 19.99Å^2. Five amino acid residues are located in generously allowed regions of the Ramachandran plot – Tyr106 from each protomer and Asp221 from one of the four protein molecules. Several side-chains were observed to have potential alternative conformations (Met95, Asp161, Leu180 and His234) but not in all of the protomers.

Table 3.3 - Crystallographic data processing and refinement statistics

	nECL	recECL	recECL
DATA PROCESSING:			
Space group	$P6_5$	P1	$P2_1$
Cell dimensions (Å)	a = 134.02	a = 55.28	a = 54.90
	b = 134.02	b = 55.37	b = 167.23
	c = 81.64	c = 86.93	c = 55.13
	$\alpha = \beta = 90°$	$\alpha = 86.23°$	$\alpha = \gamma = 90°$
	$\gamma = 120°$	$\beta = 75.37°$	$\beta = 97.09°$
		$\gamma = 82.13°$	
Resolution (Å)	2.00	2.13	1.70
Reflections measured	414,754	399,699	1,021,770
Unique reflections	56,451	55,200	108,695
Completeness (%)	99.7	93.1	94.1
in the outermost shell	99.9	86.1	77.3
$I/(\sigma I)$	18.44	10.93	28.05
in the outermost shell	6.22	5.29	5.26
R_{symm}	0.070	0.077	0.058
REFINEMENT:			
R_{cryst} (%)	18.97	19.10	17.79
R_{free} (%)	20.92	24.10	20.31
Average B factor ($Å^2$)	28.84	24.52	19.99
Wilson	23.1	25.60	19.20
RMSD from ideality:			
Bonds (Å)	0.005	0.009	0.006
Bond angles (deg)	1.44	1.47	1.49
Dihedrals (deg)	25.32	26.15	26.10
Impropers (deg)	0.78	0.91	0.78

$R_{symm} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I\rangle(hkl)|/\Sigma_{hkl}\Sigma_i I_i(hkl)$ where $\langle I\rangle$ is the averaged intensity of the *i* observations of reflection *hkl*.

$R_{cryst} = \Sigma||F_O| - |F_C||/\Sigma|F_O|$ where F_O and F_C are the observed and calculated structure factor amplitudes, respectively.

R_{free} is equal to R_{cryst} for a randomly selected 5% subset of reflections not used in the refinement.

N-linked oligosaccharide

The *N*-linked glycosylation is observed at Asn113 of the *n*ECL. The C1 atom of NAG951 forms a covalent bond (1.45Å) with ND2 atom of Asn113. The carbohydrate moiety has direct interaction with 16 water molecules. The stereo image of the heptasaccharide bound to the *n*ECL is shown in the Fig 3.4 and the electron density for the same is shown in the Fig 3.6.

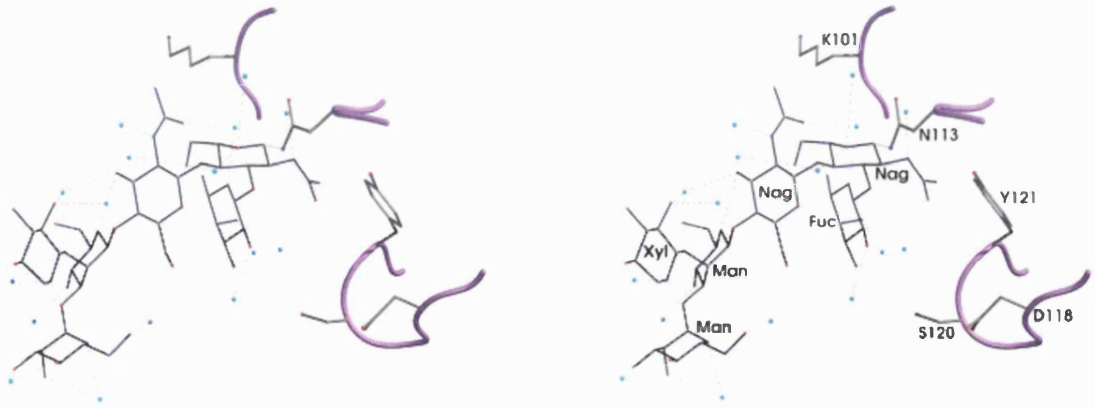


Fig 3.4 Details of the surrounding residues for the hexasaccharide at the glycosylation site (Asn113). In the crystal lattice, the N-linked carbohydrate makes indirect interaction with the combining site through water-mediated interaction. Water molecule is shown as small blue spheres.

Lactose binding site

The lactose molecule is bound to the *n*ECL by means of hydrogen bonds. The lactose forms six hydrogen bonds with residues Asn113, Gly103, Asp89, Ala218, Glu219 and has direct interaction with nine water molecules. Fig 3.5 shows the stereo image of the lactose interaction with the molecule in the asymmetric unit. Fig 3.7 shows the well defined electron density for the lactose molecule.

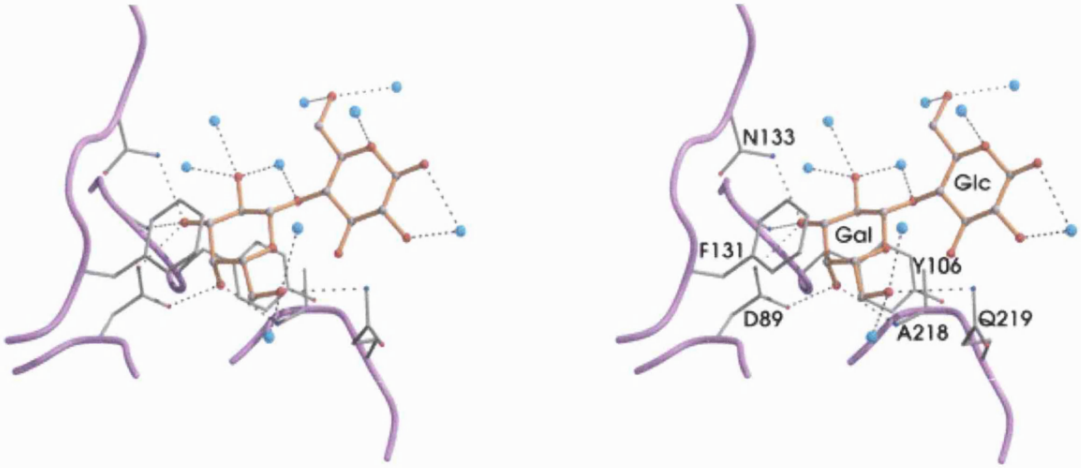


Fig 3.5 Binding of lactose to ECL. Residues involved in carbohydrate binding are: Leu86, Asp89, Gly107, Asn133, Ala218 and Gln219. Water molecules are shown as small blue spheres.

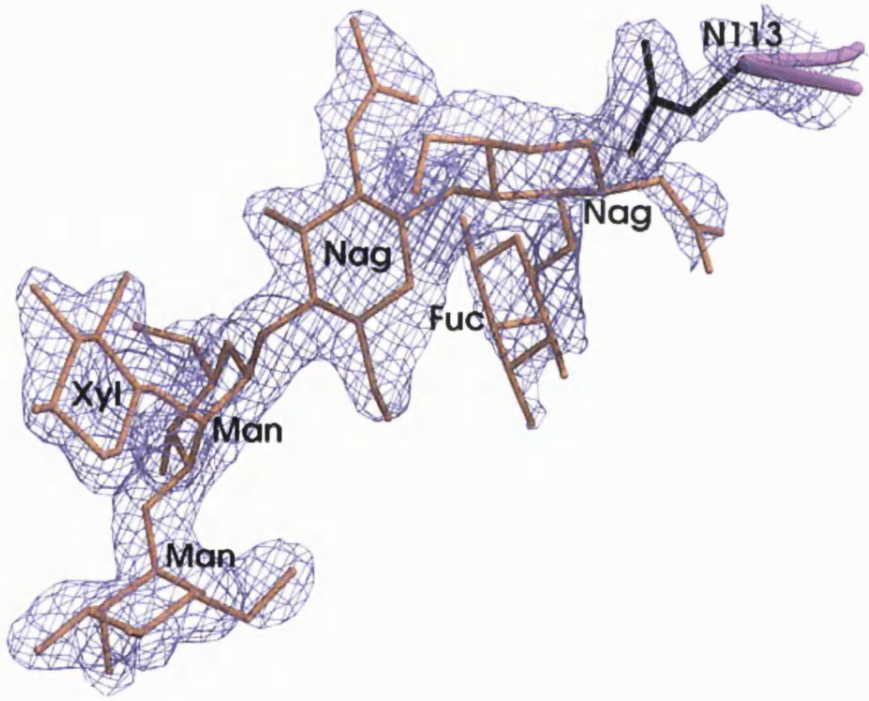


Fig 3.6 Electron density for the *N*-linked heptasaccharide (There is not enough density to build the seventh sugar unit). $2F_O$-F_C electron density map contoured at 1.0 σ.

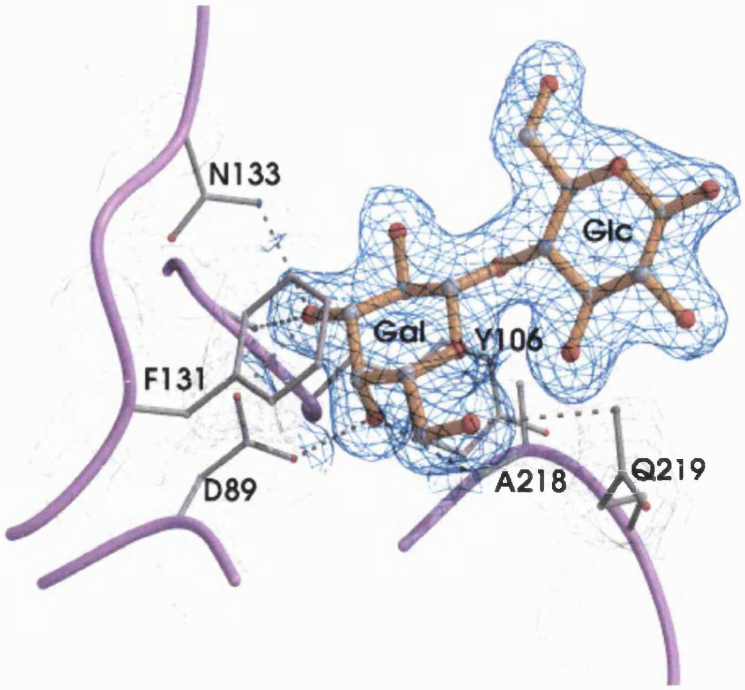


Fig 3.7 Electron density (blue) showing the disaccharide molecule and for the backbone and residues (light grey) surrounding the disaccharide. $2F_O$-F_C electron density map contoured at 1.0 σ.

The native ECL (*n*ECL) superimposes with EcorL with an overall rmsd (root mean square deviation) of 0.37Å for *Cα* atoms (238) and 0.39Å for main chain atoms (952) (Fig 3.8). The *n*ECL superimpose with recombinant ECL (*rec*ECL) with an rmsd of 0.26Å for *Cα* atoms (238).

Comparison of the two structural models indicated that there are no significant differences between the crystal structures of native and recombinant forms of ECL and, thus, confirming that *rec*ECL adopts native-like structure (rmsd = 0.26Å). The *N*-linked oligosaccharide bound to *n*ECL is covalently bound to Asn113, which is part of the loop structure located between strands β5 and β6. This residue is the only possible site of attachment for *N*-linked glycosylation in the 241-amino acid sequence of ECL. In one of the two protein molecules in the asymmetric unit, there was enough electron density to model six of the seven sugar residues bound to Asn113 (Fig 3.6). The modelled hexasaccharide has the profile: αD-Man-(1→3)-[β-D-Xyl-(1→2)]-β-

D-Man-β-D-GlcNAc-(1→4)-[α-L-Fuc(1→3)]-D-GlcNAc and does not make contacts with any other parts of the lectin protomer to which it is bound, although it might communicate with the lactose moiety bound in the combining site of a symmetry-related molecule through water mediated interactions.

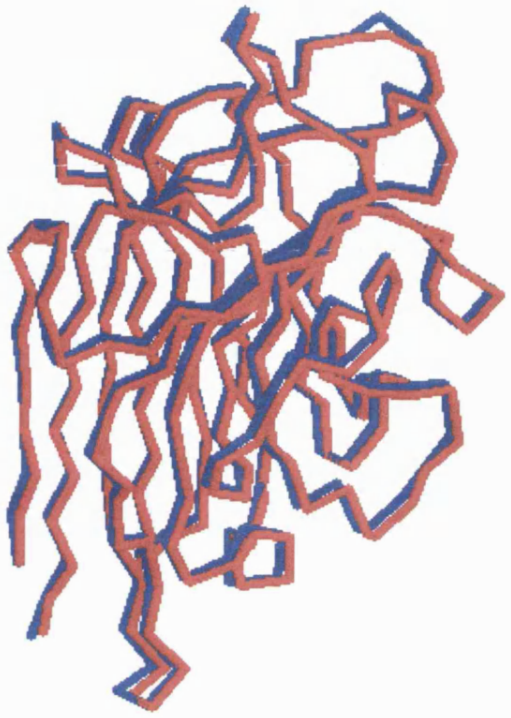


Fig 3.8 Superposition of monomers of *n*ECL (1UZY) (native ECL in blue) with EcorL (1LTE in red) for Cα-atoms (rmsd=0.37Å).

Influence of glycosylation on the quaternary structure

In *n*ECL the oligosaccharide is bound covalently to Asn113 (present study) and not to Asn17 (Shaanan et al. 1991). It was believed that the binding of the heptasaccharide to Asn17 prevents the formation of the canonical dimer (Shaanan et al. 1991). But, the *rec*ECL, which is unglycosylated at positions Asn17 or Asn113, also forms dimers back-to-back via the "handshake" motif. This mode of oligomerisation seems to be energetically more favourable than any other quaternary forms (Prabu et al. 1999; Elgavish and Shaanan 2001; Manoj and Suguna 2001; Srinivas et al. 2001). Thus, the presence or absence of *N*-linked glycosylation do not influence the mode of dimerisation in

ECL and suggests that factors intrinsic to the primary structure of the lectin dictate its quaternary structure.

Conclusion

The observation on *N*-glycosylation of ECL by Shaanan et al (1991), Svensson et al (2002) and the present analysis tells that the *N*-glycosylation is independent of the dimer formation. The lectins carbohydrate specificity and the inferences invite the researchers to plan mutational studies on ECL, making it to target different cells which will be more promising in the design of pharmaceuticals.

Chapter 4

Crystal Structure of Bovine Pancreatic Ribonuclease A in Complex with Non-natural *3'*-nucleotides

Introduction

Proteins are the final products of transcriptome of an organism and play a variety of role in a living system. Proteins that catalyse the biochemical reactions such as breakdown or make-up of bio-products in an organism are classified as enzymes. Enzymes are named after the function they perform with a suffix *-ase* in their name. The catalytic rate of reaction of an enzyme is several fold orders of magnitude than in the absence of it. Enzymes are very specific to the reactions they catalyse. Because of their specificity to the substrate and higher activity, they are molecules of interest to industry, drug companies, and biological research.

Ribonucleases (RNases) are special class of enzymes that are RNA specific and possess the ribonucleolytic activity. Ribonucleases catalyse the process of breakdown of any free RNA. There are many different ribonucleases classified and grouped as *endo-* and *exo-* ribonucleases depending on their mechanism of function. RNases play a variety of roles in a biological system and are molecules of interest in current research because of special clinical properties like cytotoxicity. Cytotoxicity of RNases depends on their catalytic activity, which is cytosolic RNA degradation thereby damping the protein production factory. This cytotoxic property of RNases has made them a molecule of interest for use in cancer therapeutics. Retargeting RNases in fusion with cell binding proteins to invade tumour cells have been reported in many researched papers (e.g. angiogenin in fusion with anti-CD22 (Newton et al. 2001) / CD30 (Huhn et al. 2001) / Epidermal Growth Factor (EGF) (Yoon et al. 1999), RNase A in fusion with EGF (Jinno et al. 1996)).

4.1 Bovine Pancreatic Ribonuclease A

Bovine pancreatic ribonuclease A (RNase A) is the first secretory endo-ribonuclease to be identified and isolated from bovine pancreas (Kunitz 1940). RNase A is relatively a small molecule of 124 amino acids and of molecular weight 14 kDa. It is a molecule of interest to biochemists and structural biologists for many years because of its remarkable biochemical and biophysical properties, for example four disulphide bonds (Fig 4.1a), which is unusual for a molecule of such a small size. RNase A catalyses the degradation of RNA by cleaving the phosphodiester bonds in RNA in the *3'* - *5'* direction resulting in *2', 3'* – cyclic phosphates (Fig 4.1b) (D'Alessio and Riordan 1997; Raines 1998).

RNase A has been well characterised in terms of its stability, structure (Wyckoff et al. 1967a; Wyckoff et al. 1967b; Wyckoff et al. 1970; Richards et al. 1972), mechanism of action, protein folding, etc., and is used as a model system by biologists because of its ease of availability (the cow pancreas) and purification. RNase A is a '*Nobel Molecule*' for it being Stanford Moore, William Stein and Christian Anfinsen shared the Nobel Prize in chemistry in the year 1972 for their work relating the chemical structure and catalytic activity of the enzyme's active site. In 1984, Bruce Merrifield was awarded the Nobel Prize for the development of novel method of chemical synthesis on a solid matrix. Bruce Merrifield used RNase A as a model system like others. Ribonucleases are molecules of scientific interest in the past few decades and now, for their properties other than the degradation of RNA, like cytotoxicity, anti-viral and anti-bacterial agents, angiogenic property, etc.

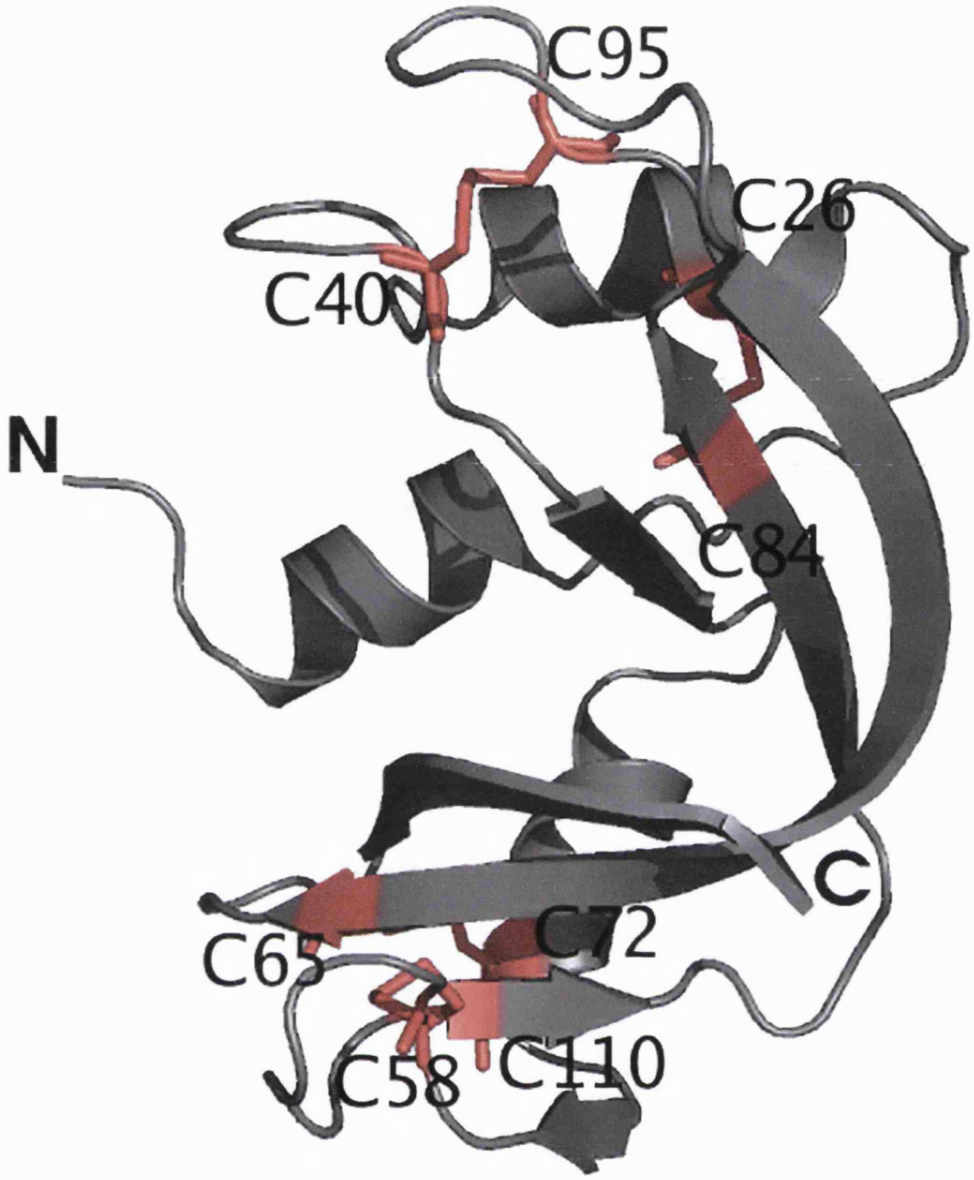


Fig 4.1a Three-dimensional structure of Ribonuclease A (PDB: 1W4O) (*grey*) and disulphide bonds (*red*) are shown. Residues involved in disulphide bridges are C40-C95; C26-C84; C58-C110; C65-C72.

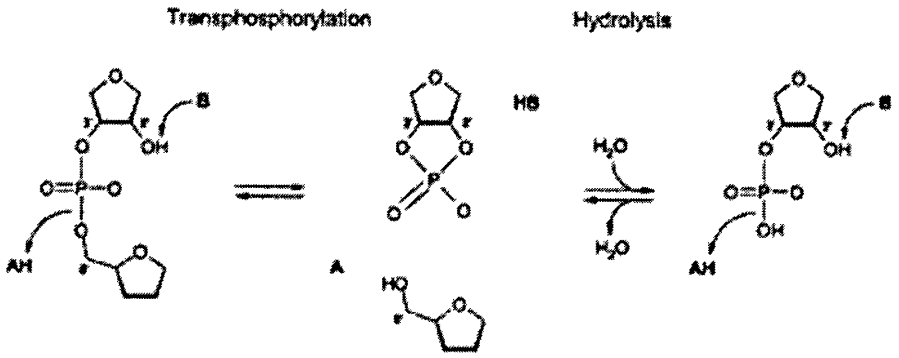


Fig 4.1b Schematic representation of formation of *2'-3'* cyclic phosphate. AH and B stands for acid and base respectively. Bent arrows shows the movement of electrons.

4.1.1 Active Site of RNase A

Structural features of RNase A

RNase A is a kidney shaped molecule (Fig 4.1a) with three α-helices, seven β-strands stabilised by four disulphide bonds - two at each pole of the molecule. RNase A is pyrimidine specific and catalyses the cleavage of phosphodiester bonds in RNA molecules. The active site of RNase A constitutes the phosphate binding sites P0, P1 & P2, and the base binding sites B1 & B2 (Fig 4.2a). A schematic representation (Pares et al. 1991) of the active site is given on the right side of Fig 4.2a. The P0 site consist only one residue K66. The P1 site, which is the primary active site, consists of residues Q11, H12, K41 and H119. H12 and H119 act as acid-base catalysts in breakdown of the phosphodiester bonds (Fig 4.1b) and form a catalytic triad with K41. The P2 site consists of residue K7. The B1 site where the pyrimidine base binds consists of residues T45, F120 and S123. The B2 site consists of residues N67, Q69, N71 and E111.

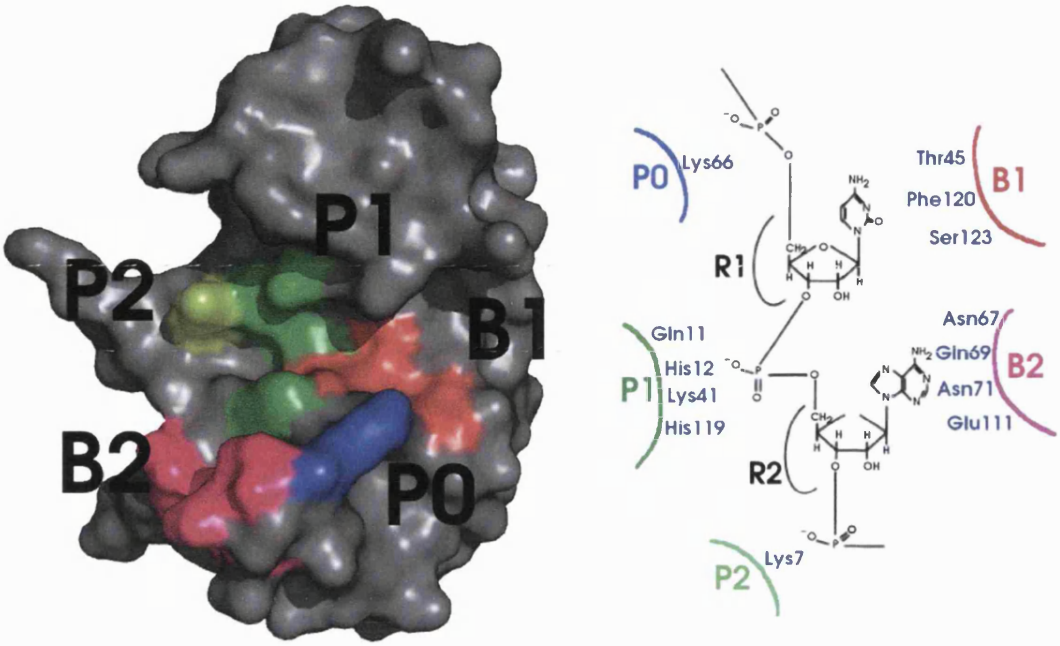


Fig 4.2a Catalytic sites of RNase A: The phosphate binding sites (P0, P1, P2) and the base binding sites (B1, B2) are represented in different colours. R1, R2 are the ribose units.

The RNA binds between the active sites in the *5'-3'* direction and the phosphate backbone is cleaved in the *3'-5'* direction between the sugars R1 and R2 (Fig 4.2a). The active site is positively charged (Fig 4.2b) thus favouring the negatively charged RNA to bind strongly. Several mutation works on the active site residues confirmed the importance of residues (Smith and Raines 2006) H12, T45 (Jenkins et al. 2005), K41 and H119 involved in the catalytic activity of the enzyme. The P0 and P2 sites in addition to the P1 site also binds the phosphates of the bases and plays a very important role in the substrate specificity and catalysis (McPherson et al. 1986; Birdsall and McPherson 1992; Fontecilla-Camps et al. 1994; Fisher et al. 1998a; Fisher et al. 1998b; Nogues et al. 1998; Cuchillo et al. 2002).

Fig 4.2b GRASP surface charge representation of ribonuclease A showing that the active site is highly concentrated with basic residues resulting in high positive charge. Picture adapted from Ribonucleases, Part A, *Methods in Enzymology*, 341, pp-194.

4.1.2 RNase A Superfamily

Ribonuclease A and its homologues together constitute Ribonuclease A superfamily, so far reported only in vertebrates (Beintema and Neuteboom 1983; Beintema et al. 1988; Beintema 1998; Rosenberg 1998; Zhang et al. 2000; Rosenberg et al. 2001; Cho et al. 2005; Dyer and Rosenberg 2006). Ribonucleases were also reported in prokaryotes (Oda et al. 1991; Katayanagi et al. 1993) and plants (Abel and Glund 1986; Konieczny et al. 1991; Taylor et al. 1993).

Fig 4.3b Multiple sequence alignment of ribonucleases of RNase A superfamily. Conserved catalytic residues are boxed in red.

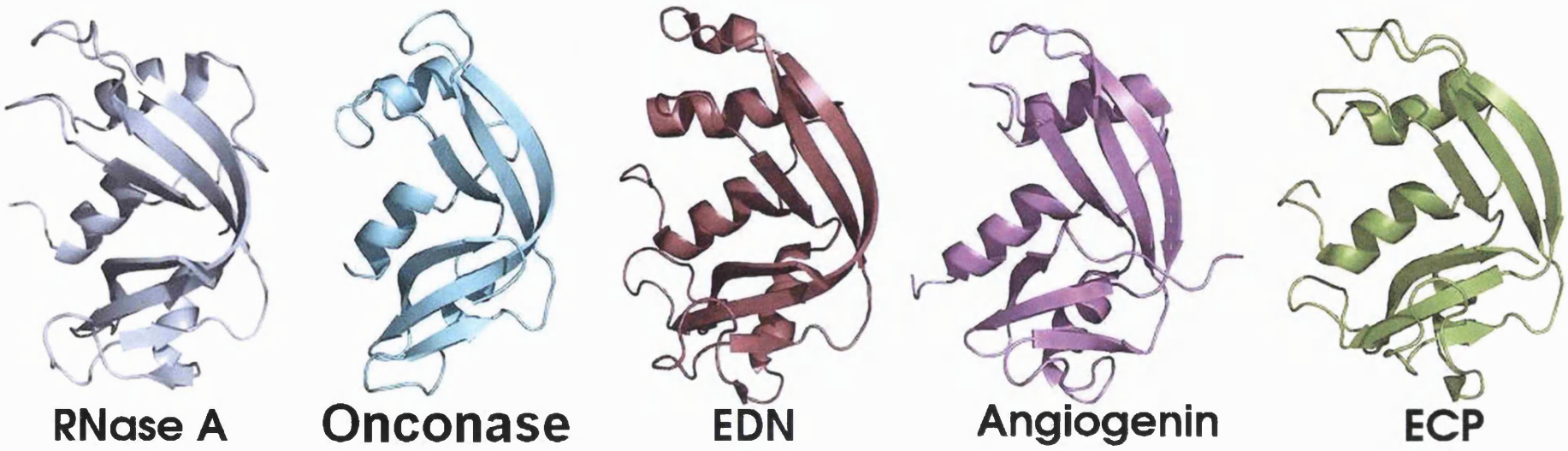


Fig 4.3a Structural similarity between RNases of RNase A superfamily. RNase A (PDB: 1W4O); Onconase (PDB: 1ONC); EDN (PDB: 1HI2); Angiogenin (PDB: 1H52); ECP (PDB: 1QMT)

RNase A superfamily consists of ribonucleases of pancreatic origin and eosinophil origin. The ribonucleases belonging to RNase A superfamily share similar fold/topology in their tertiary structure, fold architecture and overall sequence similarity, and identity at their catalytic site residues (Fig 4.3a,b). Some of RNase A homologues are Angiogenin (RNase 5), Eosinophil-Derived Neurotoxin (EDN/RNase 2), Eosinophil Cationic Protein (ECP/RNase 3), Onconase (Frog ribonuclease), and RNase 4. RNase A shares an overall sequence similarity of 44%, 46%, 54%, and 46% with onconase, EDN, angiogenin, and ECP respectively.

Brief description of RNase A homologues

Angiogenin: A ~14 kDa protein and a member of RNase A superfamily with unusual biological properties (Strydom et al. 1985; Bicknell and Vallee 1988; 1989; Bond et al. 1993; Jimi et al. 1995; Soncin et al. 1997; Hu et al. 2000) and exhibit weak ribonucleolytic activity (Shapiro et al. 1986; St Clair et al. 1988). Angiogenin shares an overall sequence identity of 40% to RNase A (Strydom et al. 1985) and involved in blood vessel formation and tumour cell growth. Angiogenin is found in different species including human, mouse, bovine, zebrafish (Pizzo et al. 2006), and many others. High-level expression of recombinant human angiogenin and murine angiogenin in *Escherichia coli* has been reported (Holloway et al. 2001). Structural studies on bovine, human and murine angiogenin were also reported highlighting the catalytic residues involved in substrate binding, key residues mediating cell surface binding of angiogenin, and residues involved in molecular recognition (Shapiro et al. 1989; Shapiro and Vallee 1989; Acharya et al. 1992; Acharya et al. 1994; Russo et al. 1994; Acharya et al. 1995; Leonidas et al. 1999a; Jardine et al. 2001; Leonidas et al. 2001b; Holloway et al. 2004; Holloway et al. 2005).

ECP: Eosinophil Cationic Protein is ~15.5 kDa in molecular weight and found in eosinophil granules. It shares an overall sequence identity of 33% with RNase A. Structural characterisation of ECP revealed the formation of catalytic triad similar to RNase A (Boix et al. 1999) and thus ECP possesses ribonucleolytic activity which is responsible for its functional properties. *In vitro* studies of ECP has shown that they function as anti-bacterial (Lehrer et al. 1989) and anti-viral (Harrison et al. 1999) agents. Neurotoxic effect of ECP results in *Gordon phenomenon*, a muscular paralysis. Other properties of ECP are helminthotoxicity (Hamann et al. 1990), and cytotoxicity (Young et al. 1986). Reports have shown that ECP might be involved in degranulation of eosinophils (Klion and Nutman 2004; Lee and Lee 2005).

EDN: Eosinophil-Derived Neurotoxin is a major constituent of human eosinophil leukocytes sharing an overall sequence identity of 34% to RNase A (Durack et al. 1981; Gleich et al. 1986) and ~70% sequence similarity to ECP weighing about 18.4 kDa. Similar to ECP EDN also possesses neurotoxicity effect leading to the loss of Purkinje cells in cerebellum. The catalytic centre of EDN is structurally characterised and reported confirming that ribonucleolytic property is important for its neurotoxicity (Leonidas et al. 2001a). Recent investigations have shown that EDN acts as a chemotactic agent in dendritic cells (Yang et al. 2003; Oppenheim and Yang 2005).

Onconase: Onconase is an amphibian ribonuclease from *Rana pipiens* (Ardelt et al. 1991). It shares an overall sequence identity of 32% to RNase A. The protein possesses anti-tumour activity, which makes it a molecule of clinical/therapeutic interest and is in phase II/III clinical trials in the treatment of unresectable malignant mesothelioma (UMM) and Non-Small Cell Lung Cancer (NSCLC) (http://www.alfacell.com/).

4.2 Inhibitors of Ribonucleases

The digestion of cellular RNA by RNases leads to RNA degradation. If this process is allowed to continue there will be no protein production in the cell and all RNases might become cytotoxic resulting in cell death (Vescia et al. 1980; Wu et al. 1993; Leland et al. 1998; Suzuki et al. 1999). Nature has found a way to inhibit the ribonucleolytic activity of the RNases. Similar to RNases capturing free RNA molecules, ribonuclease inhibitors captures free RNases and makes them inactive.

4.2.1 Natural Inhibitor of Ribonucleases – Ribonuclease Inhibitor (RI)

The natural ribonuclease inhibitor (RI) is ~50 kDa molecular weight Leucine–rich repeat (LRR) protein. RI is a horseshoe shaped molecule (Fig 4.4) inner and outer diameter of ~21 Å and ~67 Å respectively. It consists of 16 β-strands and 16 α-helices connected by β/α and α/β loops. It exhibits very high affinity in the order of fM (*femto molar*) (Bicknell and Vallee 1989; Shapiro and Vallee 1991; Bond et al. 1993; Hofsteenge et al. 1998) range to RNases of RNase A superfamily, exception includes onconase where the affinity is of the order of >1 µM (*micro molar*) (Boix et al. 1996). RI has been isolated and purified from human (hRI), rat (rRI), and porcine (pRI) sources. Crystal structures of RI on its own (Kobe and Deisenhofer 1993) and in complex with RNase A (Kobe and Deisenhofer 1995; 1996), angiogenin (Papageorgiou et al. 1997), and EDN (Iyer et al. 2005) have been reported.

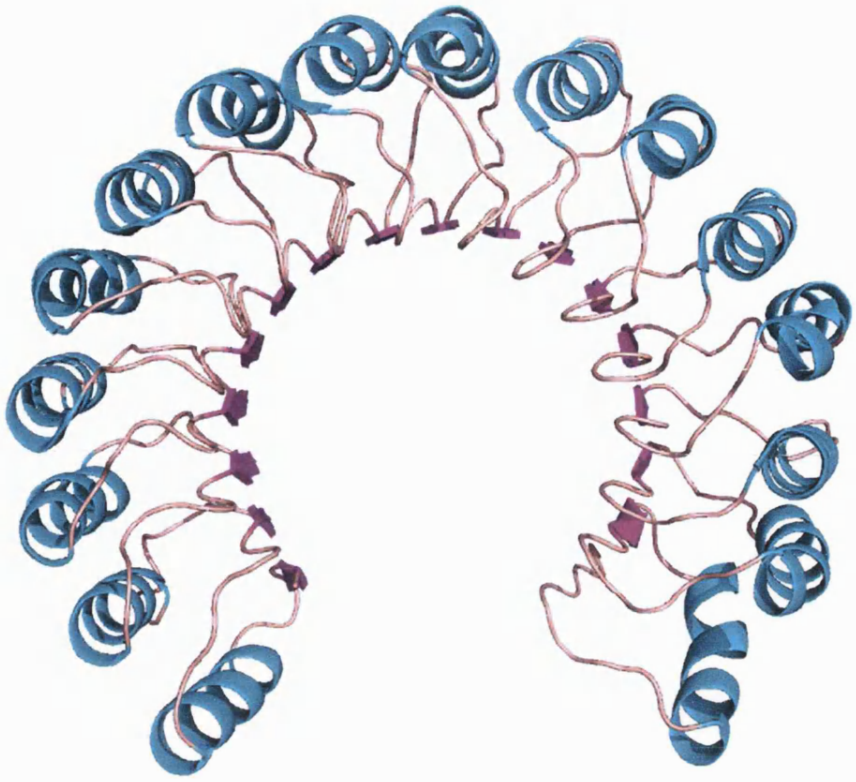


Fig 4.4 Ribbon diagram representation of Ribonuclease Inhibitor (RI) (PDB: 2BNH). (Kobe and Deisenhofer 1993)

4.2.2 Small Molecule Inhibitors of Ribonuclease A

In order to understand the mechanism of action of an enzyme one needs to understand the chemical nature and structure of the active site. Initial studies were done by subjecting/treating the enzyme under suitable chemical conditions (Weil and Seibles 1955; Barnard and Stein 1959; Gundlach et al. 1959). Small molecule inhibitor studies on RNase A as a model system will aid us in understanding the mechanism of function of the enzyme in living organism and the key residues involved in ribonucleolytic activity. Most of the RNase A inhibitors are potent inhibitors of EDN, Angiogenin, and RNase 4 (Russo and Shapiro 1999; Russo et al. 2001). Thus an effective inhibitor of RNase A could be tested on its homologues.

An inhibitor could be competitive or non-competitive. Competitive inhibition depends solely on the affinity of the inhibitor to the enzyme's active site resulting in an inactive enzyme-inhibitor complex. The strength of inhibition is defined by K_i values, which is a quantitative measure of inhibitor strength. Small molecule binding studies to RNase A has revealed the important catalytic residues of the enzyme and also that these molecules could be used as inhibitors of RNase A activity. Some of the small molecule inhibitors are bromoacetate (Barnard and Stein 1959), iodoacetate (Gundlach et al. 1959), 2'-deoxyuridine 3'-phosphate (Walz 1971), and arabinonucleotides (Pollard and Nagyvary 1973).

Several small molecule inhibitors of RNase A have been reported from our laboratory. A few of them are listed in the Table 4.1.

Table 4.1

Inhibitor	K_i	References
ppA-3'-p	~0.24 mM	(Leonidas et al. 1997; Russo et al. 1997)
ppA-2'-p	~0.52 mM	(Leonidas et al. 1997; Russo et al. 1997)
dUppA-3'-p	~11.3 µM	(Jardine et al. 2001)
pdUppA-3'-p	~0.027 µM	(Leonidas et al. 1999b; Russo and Shapiro 1999)

4.3 RNase A in Complex with Non-natural *3'*-nucleotides

RNA based inhibitor design for RNase A has been the most common method used (Russo et al. 1997; Russo and Shapiro 1999; Russo et al. 2001). Any approach had the ribose/deoxy-ribose sugar unit as a standard (Stowell et al. 1995).

Does the ribose/deoxy-ribose molecule play an important role in binding of the inhibitor to RNase A?

Here, we report the synthesis and structural details of the complexes with RNase A of two 3'-nucleotides containing nonnatural furanose rings: 2'-fluoro-2'-deoxyuridine 3'-phosphate and arabinouridine 3'-phosphate.

4.3.1 MATERIALS AND METHODS

4.3.1.1Synthesis of Non-natural 3'-nucleotides

The synthesis of non-natural 3'-nucleotides and kinetic studies were done in our collaborator Professor Ronald Raines Laboratory, University of Wisconsin-Madison, USA, and reported along with the structural complex of RNase A (Jenkins et al. 2005).

4.3.1.2 Crystallisation, Data Collection and Structure Solution

Bovine pancreatic RNase A was obtained from Sigma. RNase A crystals were grown using the vapour diffusion technique as described in ref (Leonidas et al. 1997); they belong to the space group C2, with two molecules per asymmetric

unit. Crystals of the complexes were obtained by soaking the RNase A crystals in 20 mM sodium citrate, pH 5.5, and 25% PEG 4000, 50 mM dUMP, 1 mM araUMP, 12.5 mM $dU^{F}MP$ for 45, 60 and 75 minutes respectively prior to data collection. Diffraction data for the three complexes were collected at 100 K (the reservoir buffer with 30% PEG 4000 was used as cryoprotectant) on stations PX 14.1 and PX 9.6 from a single crystal at the Synchrotron Radiation Source, Daresbury, UK using an ADSC Quantum 4 CCD detector. All diffraction images were integrated using HKL2000 (Otwinowski and Minor 1997b). Initial phases were obtained using the structure of free RNase A (Leonidas et al. 1997) as a starting model. The refinement was carried out using the CNS suite (Brunger et al. 1998) and the model building was carried out using O (Jones et al. 1991). Initial model building and refinement was carried out without the inhibitor. In each data set, a set of reflections was kept aside for R_{free} calculation (Brunger 1992). The inhibitor and water molecules were modelled using both the 2Fo-Fc and Fo-Fc SIGMAA weighted maps. The topology and parameter files for the inhibitors were either generated manually and/or using the Hic-up server (Kleywegt and Jones 1998b) http://alpha2.bmc.uu.se/hicup/. All three structures have good geometry, and more than 87% of the protein residues are in the most favourable region of Ramachandran plot. All structural diagrams were prepared using the program Bobscript (Esnouf 1997b).

4.3.1.3 Determination of K_i values for the inhibitors

The values of K_i for the four 3'-nucleotides were measured by their ability to inhibit the cleavage of the fluorogenic substrate 6-FAM–dArU$(dA)_2$–TAMRA by wild-type RNase A and its T45G variant (Kelemen et al. 1999), and are listed in Table 4.2. All four 3'-nucleotides were potent inhibitors of the wild-type enzyme, whereas inhibition of T45G RNase A was less pronounced—by up to three orders-of-magnitude. Phosphoryl group pK_a values were determined by ^{31}P NMR spectroscopy. Inhibition was determined in MES–NaOH buffer, pH 6.0, containing NaCl (50 mM).

Table 4.2 Values of pK_a and K_i for 3'-nucleotides. Chemical structures of non-natural 3'-nucleotides are given in Fig 4.5a

Nucleoside 3'-phosphate	pK_a	K_i (μM)	
		Wild-type RNase A	T45G RNase A
3'-UMP	5.84 ± 0.05	39 ± 2	89 ± 9
dUMP	6.29 ± 0.07	18 ± 3	≥1700
dU^FMP	5.89 ± 0.10	5.5 ± 0.7	181 ± 15
araUMP	5.85 ± 0.06	6 ± 1	≥1000

4.4 X-ray crystallographic structures of dUMP, dU^FMP and araUMP complexes

The 3D structures of the three complexes of RNase A were determined at high resolution using X-ray crystallography (Table 4.3). The binding of the nucleotides at the catalytic site were clear in all three complexes (except for molecule B in araUMP complex due to severe cracking of crystals while soaking) as observed from the electron density maps (Fig 4.5b). The complex structures superimpose well with that of the RNase A native structure with an rms deviation in the range of 0.52Å. The interactions made by dUMP, dU^FMP and araUMP complexes are similar (Fig. 4.6a-c, Table 4.6a-c) and the nucleotides predominantly bind at sites P_1, B_1 and P_2. The predominant hydrogen bonds were observed between uracil and Thr45; the 3'-R-phosphate with the side chains of Gln11, His12, Lys41, and His119, and the main chain of Phe120 at the catalytic site (Fig 4.6d).

Table 4.3: Crystallographic statistics

RNase A complex	dUMP	araUMP	$dU^{F}MP$
Resolution (Å)	50-1.69	50-1.60	50-1.68
(Outermost shell) (Å)	1.75-1.69	1.66-1.60	1.74-1.68
Reflections measured	184,596	455,952	239,956
Unique reflections	27,984	30,191	27,669
R_{symm}[a]	0.093	0.086	0.047
(Outermost shell)	0.122	0.517	0.086
Completeness			
(outermost shell) (%)	91.4 (94.2)	89.9 (88.5)	94.8 (87.6)
<I/σI> (outermost shell)	9.83	10.3	5.18
R_{cryst}[b]	0.22	0.23	0.21
(Outermost shell)	0.21	0.50	0.29
R_{free}[c]	0.23	0.25	0.24
(Outermost shell)	0.27	0.54	0.33
Number of solvent molecules	291	203	310
RMS deviation from ideality			
in bond lengths (Å)	0.010	0.005	0.004
in angles (°)	1.5	1.2	1.3
Average B factor ($Å^2$)	18.4	27.92	14.95

[a] $R_{symm} = \Sigma_h\Sigma_i|I(h) - I_i(h)/\Sigma_h\Sigma_i I_i(h)$, where, $I_i(h)$ and $I(h)$ are the i^{th} and the mean measurements of the intensity of reflection *h*.

[b] $R_{cryst} = \Sigma_h|F_o - F_c|/\Sigma_h F_o$, where, F_o and F_c are the observed and calculated structure factors amplitudes of reflection *h*, respectively.

[c] R_{free} is equal to R_{cryst} for a randomly selected 5% subset of reflections not used in the refinement (Brunger 1992).

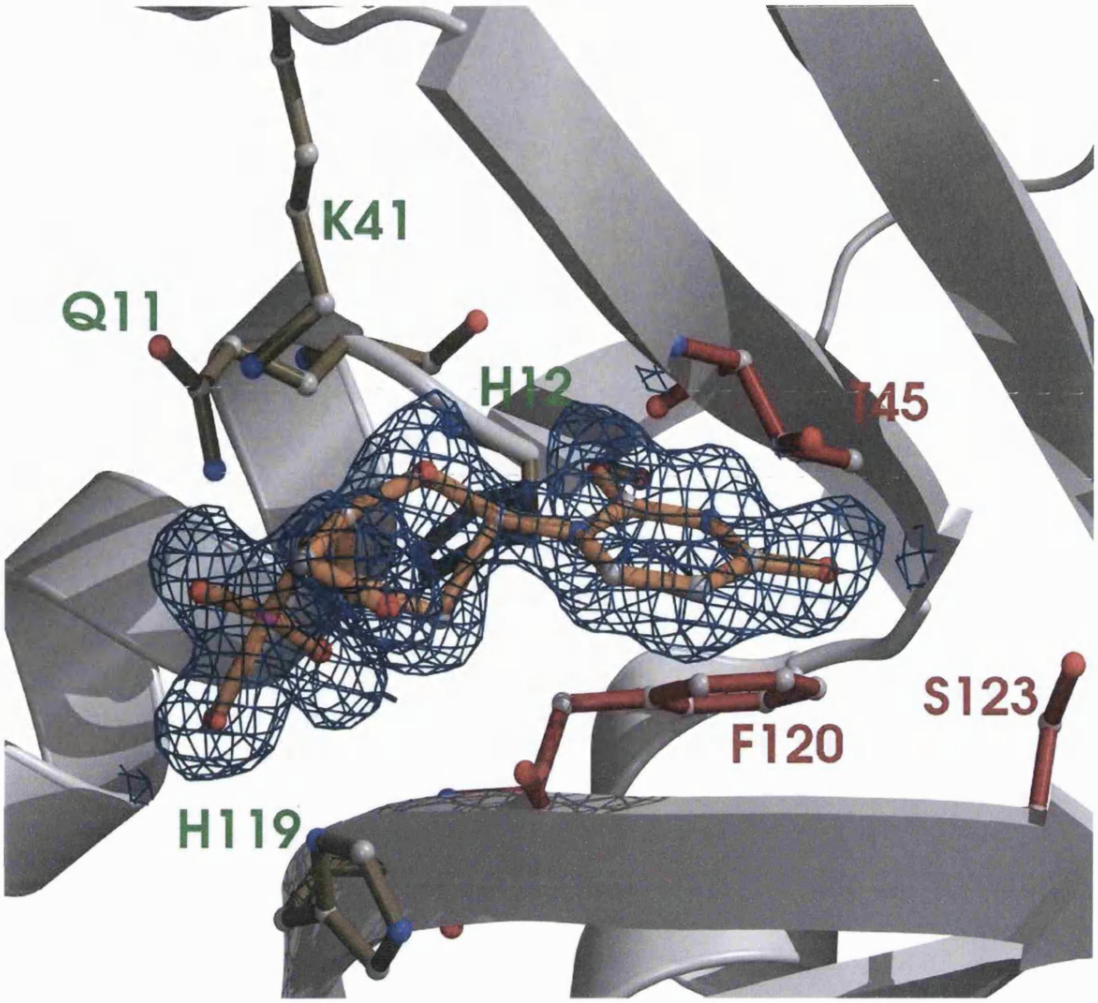


Fig 4.6d A ribbon diagram and ball-and-stick representation of active sites (P1, B1) of RNase A·dUMP complex. The dUMP is shown as ball-and-stick (gold) model with the well-defined electron density around it. $2F_O$-F_C electron density map contoured at 1.0 σ.

In all three complexes the uridine binds in an anti conformation, and its ribose adopts a C3'-exo, C2'-endo and O4'-endo pucker in dUMP, araUMP and dU^FMP respectively (Altona and Sundaralingam 1972) (Table 4.7). The difference in conformation of the ribose moiety can be attributed to the presence of additional atom/s attached at C2' position (-OH group in araUMP and fluorine atom in the dU^FMP complex) (Fig. 4.7b, Fig 4.8a-b).

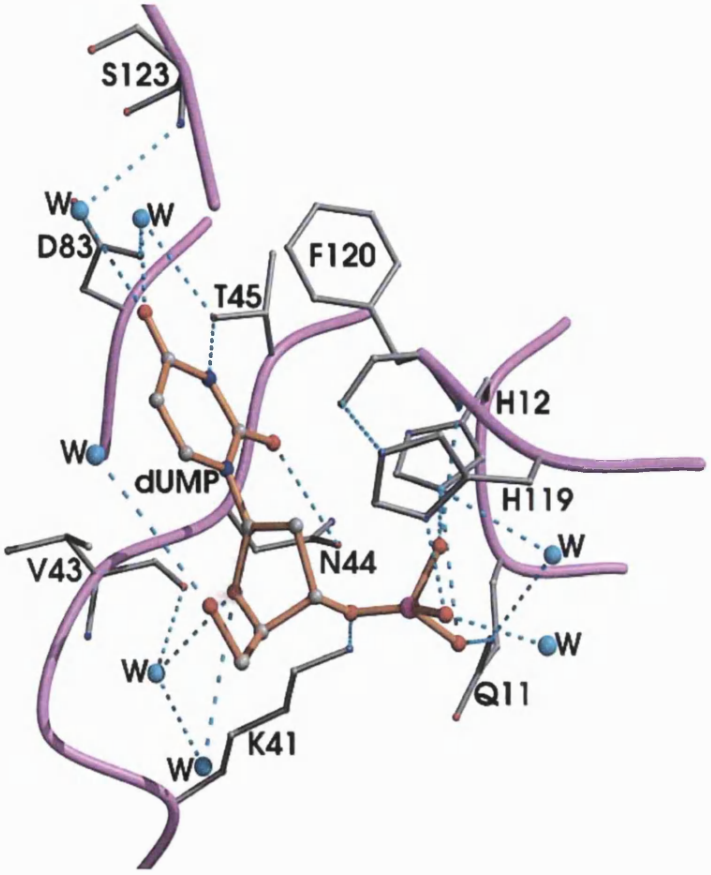


Fig 4.6a Details of the interactions at the catalytic site of RNase A molecule and dUMP. Water molecules are represented by small spheres. Hydrogen bonds are indicated by dashed lines.

Table 4.6a Potential hydrogen bonds of dUMP with RNase A

dUMP atom	RNase A residue	Distance (Å)
N3	Thr45 - OG1	2.82
O1P	His12 - NE2	2.83
O1P	Phe120 - N	3.01
O3P	His119 - ND1	2.54
O2'	His12 - NE2	2.86
O2'	Gln11 - NE2	2.83
O3'	Lys41 - NZ	3.20
O4	water	3.29
O4	water	2.74
O3P	water	2.71
O2'	water	3.29
O4'	water	3.04

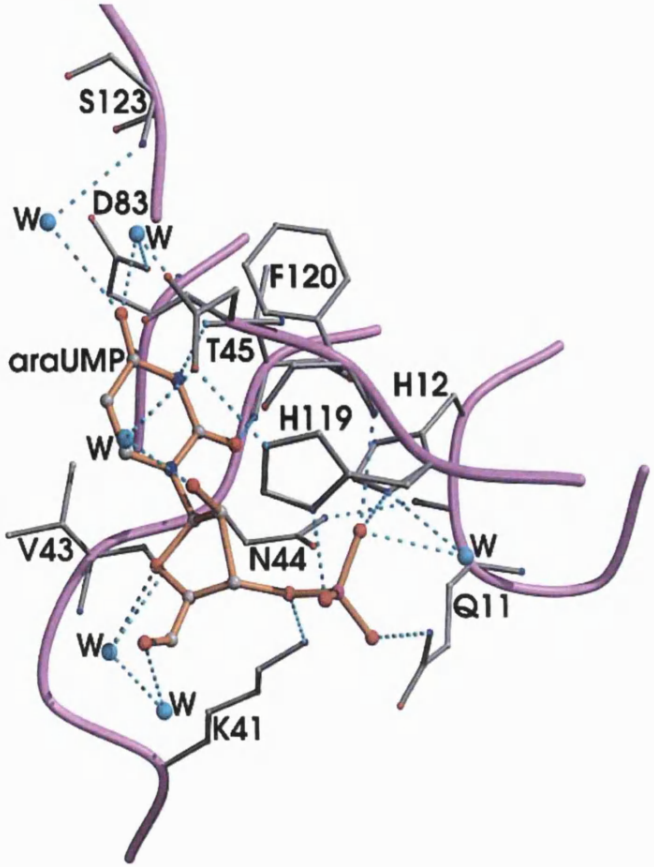


Fig 4.6b Details of the interactions at the catalytic site of RNase A molecule and araUMP. Water molecules are represented by small spheres. Hydrogen bonds are indicated by dashed lines.

Table 4.6b Potential hydrogen bonds of araUMP with RNase A

araUMP atom	RNase A residue	Distance (Å)
O2	Thr45 - N	2.87
N3	Thr45 - OG1	2.70
O3P	His12 - NE2	2.71
O3P	Phe120 - N	2.89
O1P	Gln11 - NE2	2.95
O2'	His119 - ND1	2.66
O3'	Lys41 - NZ	3.11
O4'	Val43 - O	3.34
O4	water	3.27
O4	water	2.59
O1P	water	3.19
O3P	water	2.68
O2'	water	2.84
O4'	water	2.97

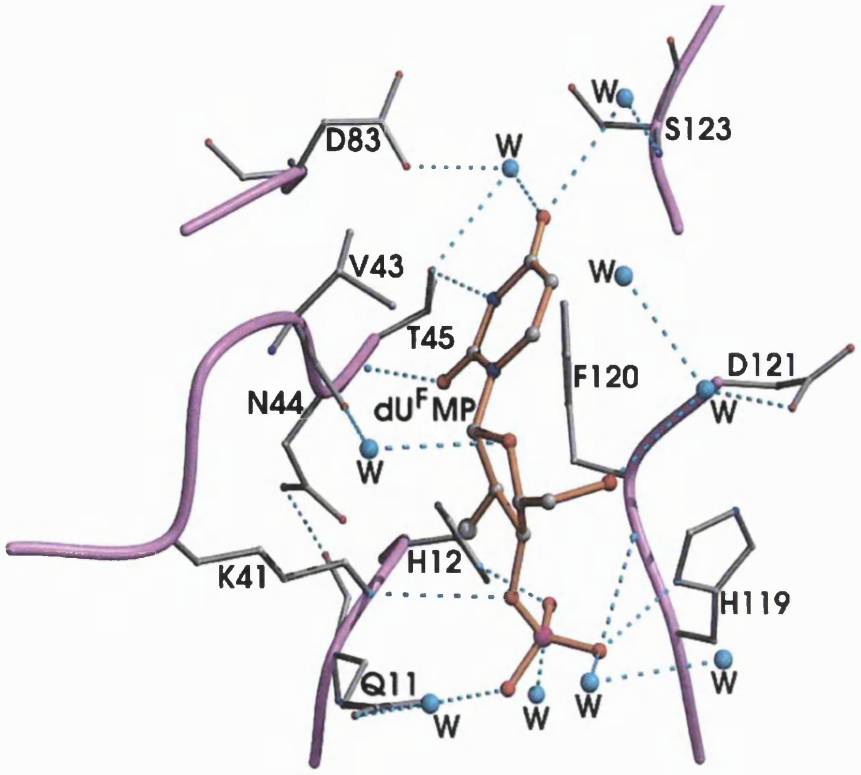


Fig 4.6c Details of the interactions at the catalytic site of RNase A molecule and $dU^{F}MP$. Water molecules are represented by small spheres. Hydrogen bonds are indicated by dashed lines.

Table 4.6c Potential hydrogen bonds of $dU^{F}MP$ with RNase A

$dU^{F}MP$ atom	**RNase A residue**	**Distance (Å)**
O2	Thr45 - N	2.61
N3	Thr45 - OG1	2.71
O1P	Gln11 - NE2	2.93
O3P	His12 - NE2	2.75
O2'	His119 - ND1	2.50
O2'	Phe120 - N	3.09
O3'	Lys41 - NZ	3.20
O4	water	3.41
O4	water	2.61
O1P	water	2.53
O3P	water	2.95
O2'	water	2.96

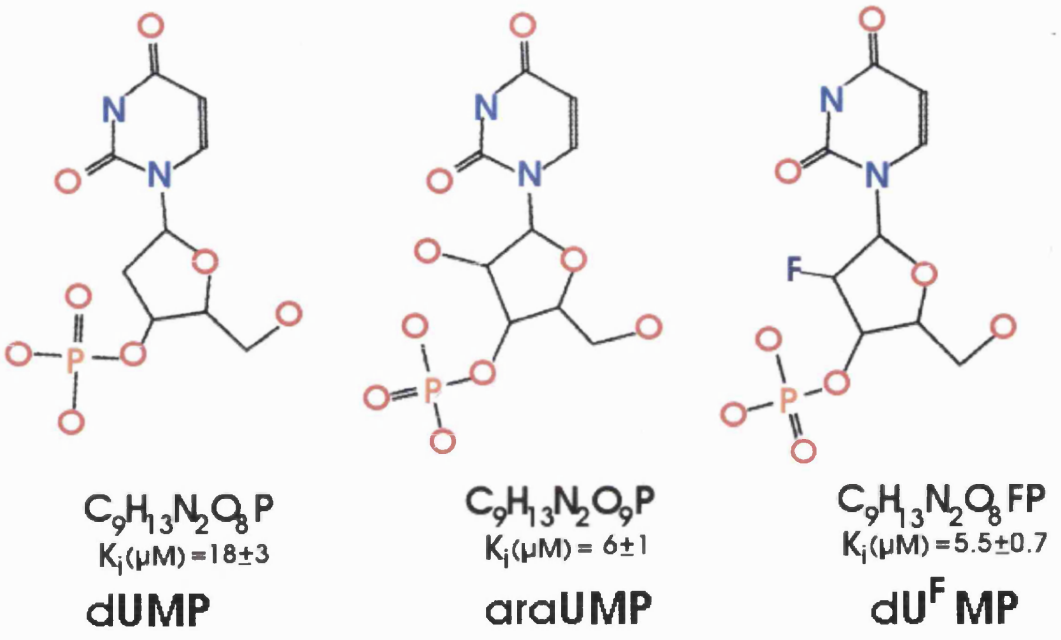


Fig 4.5a Chemical structure of non-natural 3'-nucleotides

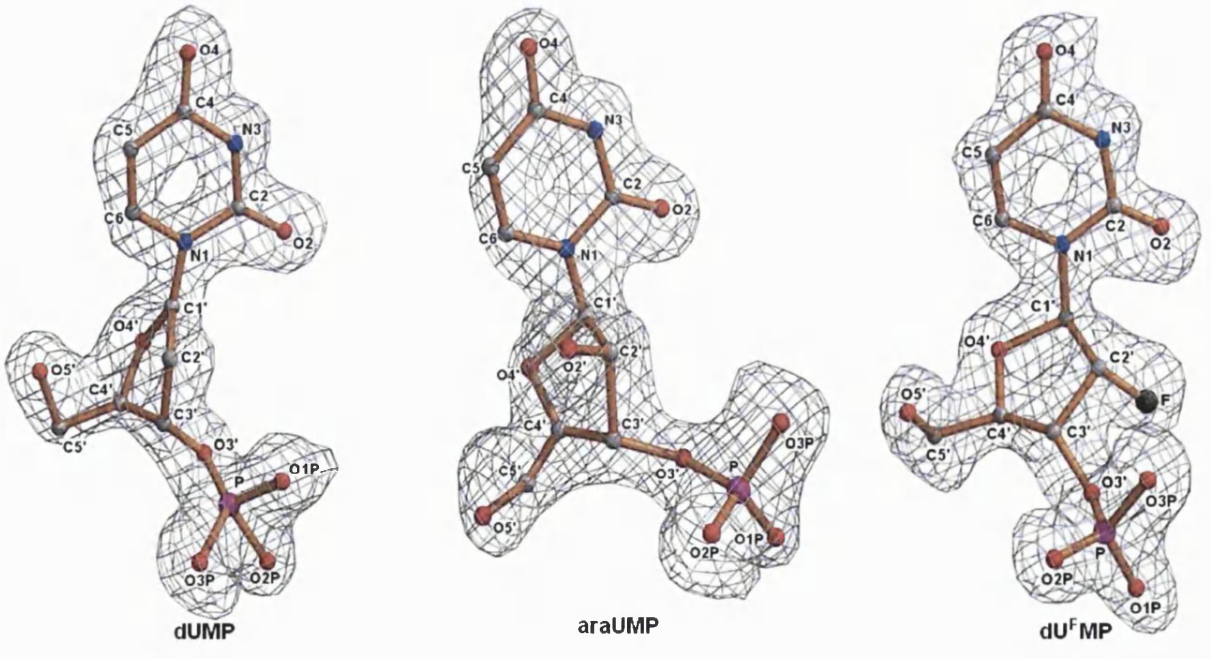


Fig 4.5b Portions of the electron density maps (2Fo-Fc) for dUMP, araUMP and dU^FMP respectively. $2F_O$-F_C electron density map contoured at 1.0 σ.

S conformation
$C_{2'}$-endo–$C_{3'}$-exo

N conformation
$C_{2'}$-exo–$C_{3'}$-endo

Fig 4.7a S and N conformation of nucleosides.

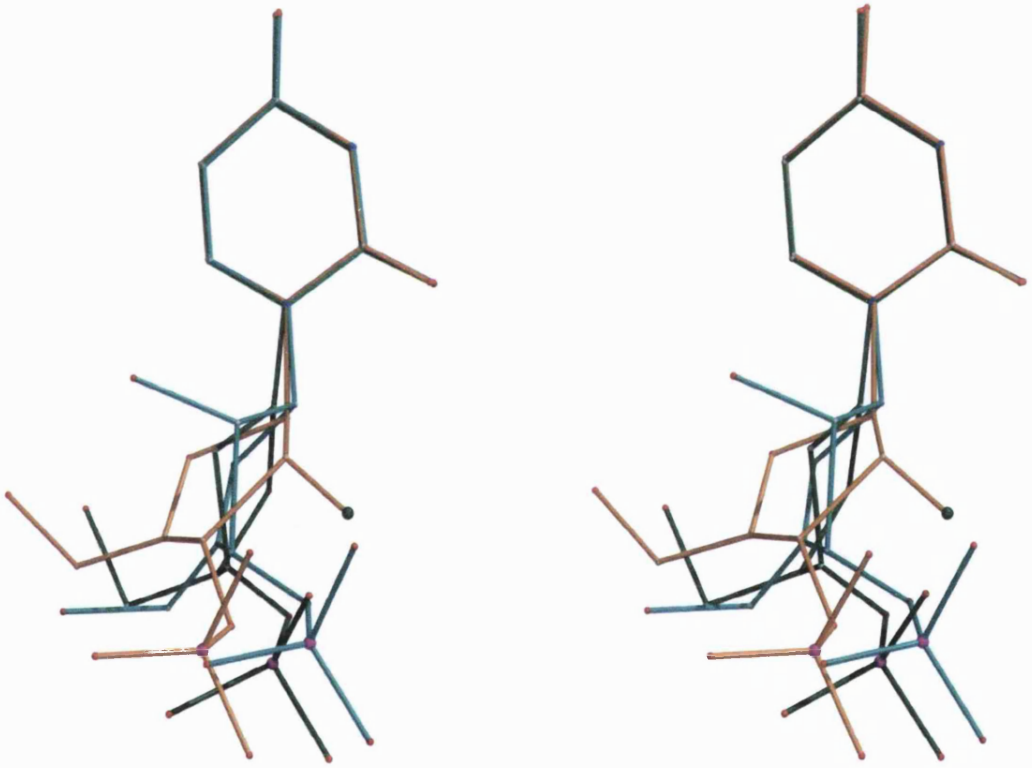


Fig 4.7b Stereo representation of dUMP-dark green; araUMP- blue; $dU^{F}MP$-gold.

Table 4.7: Torsion angles statistics for dUMP, araUMP, $dU^{F}MP$ when bound to RNase A

RNase A molecule	**dUMP**	**araUMP**	**$dU^{F}MP$**
Backbone torsion angles			
O5'-C5'-C4'-C3' (γ)	86.3 (+*sc*)	73.0 (+*sc*)	62.9 (+*sc*)
C5'-C4'-C3'-O3' (δ)	123.4 (+*ac*)	94.8 (+*ac*)	66.7 (+*sc*)
C5'-C4'-C3'-C2'	-99.7	-145.7	-158.4
C4'-C3'-C2'-O2'	--	93.0	-134.1
Glycosyl torsion angles			
O4'-C1'-N1-C2 (χ')	-107.3 (*anti*)	-156.2 (*anti*)	176.0 (*anti*)
Pseudorotation angles			
C4'-O4'-C1'-C2' (υ_0)	2.5	-40.3	-27.9
O4'-C1'-C2'-C3' (υ_1)	23.6	50.9	16.3
C1'-C2'-C3'-C4' (υ_2)	-41.8	-39.3	-2.8
C2'-C3'-C4'-O4' (υ_3)	42.7	17.9	-16.0
C3'-C4'-O4'-C1' (υ_4)	-28.9	15.2	28.8
Phase	201	144	96
	C3'-exo	*C2'-endo*	*O4'-endo*
C2'-C3'-O3'-P	85.1	89.2	103.5
C4'-C3'-O3'-P	-147.4 (-*ac*)	-158.4 (+*ap*)	-121.8 (-*ac*)

4.5 Results and Discussion

3'-Nucleotides with a nonnatural furanose ring can have a greater affinity for wild-type RNase A than do 3'-nucleotides with a ribose or deoxyribose ring. The K_i values increase in the order: $dU^FMP \approx$ araUMP < dUMP < 3'-UMP (Table 1). 3'-Nucleotides with a dianionic phosphoryl group is known to be more potent inhibitors of RNase A than those with a monoanionic phosphoryl group (Russo et al. 2001). The difference in pK_a values between dUMP and the other three nucleoside 3'-phosphates is approximately 0.4 pK_a units (Table 4.2). Likewise, the K_i value of dUMP with wild-type RNase A is approximately 3-fold higher than those of dU^FMP and araUMP. 3'-UMP has a higher K_i value with wild-type RNase A than would be expected based on its pK_a. This weaker binding could arise from the 2'-OH group participating in more unfavourable interactions with the enzyme than do the smaller 2'-groups (F or H) of the other 3'-nucleotides.

These unfavourable interactions could be reinforced by the tight interaction between the uracil base and Thr45, and by a slight preference for the C3'-endo (N) conformation of 3'-UMP in solution (Davies and Danyluk 1975) (Fig 4.7a). It is interesting to note that in the structure of the crystalline RNase A·3'-UMP complex (Leonidas et al. 2003), the ribose ring is found in the C2'-endo rather than the C3'-endo conformation. Like 3'-UMP, dU^FMP resides in the C3'-endo conformation (Antonov et al. 1976), but has a smaller fluoro group in its 2'-position. dUMP (Guschlbauer and Jankowski 1980) and araUMP (Chwang and Sundaralingam 1973; Venkateswarlu and Ferguson 1999) are predominantly in the C2'-endo (S) conformation and both have hydrogens where 3'-UMP has a larger hydroxyl group. The weaker binding of 3'-UMP compared to dUMP and araUMP lends supports the hypothesis that ground-state destabilization contributes to the catalytic prowess of RNase A (Kelemen et al. 2000).

Another possible reason for the K_i value of 3'-UMP being greater than those of the other 3'-nucleotides with wild-type RNase A is that the energetic penalty of desolvating the 2'-OH group is higher than for any of the other 2'-substituents. The 2' H and F of dUMP and dU^FMP are more easily desolvated than is a hydroxyl group, due to their lack of ability form hydrogen bonds with water (Howard et al. 1996), and the 2'-arabino hydroxyl group of araUMP might not need to be desolvated for that 3'-nucleotide to bind to RNase A because it is oriented away from the active-site residues.

The binding of 3'-nucleotides by T45G RNase A follows a much different trend than does binding by the wild-type enzyme. The K_i values increase in the order: 3'-UMP < dU^FMP << araUMP,dUMP. 3'-UMP has half of the affinity for T45G RNase A than for the wild-type enzyme, whereas dU^FMP has 3% of the affinity. The K_i values for dUMP and araUMP could not be measured, but were estimated to be ≥1.7 and ≥1 mM, respectively. The general decrease in binding between T45G RNase A and the inhibitors underlines the importance of the interaction between Thr45 and pyrimidine nucleobases in substrate binding (delCardayré and Raines 1994; 1995; Kelemen et al. 2000). The binding of 3'-nucleotides to T45G RNase A appears to be sensitive to the pucker of the furanose ring: 3'-UMP and dU^FMP, which both reside in the C3'-endo conformation, bind with high micromolar affinities, whereas dUMP and araUMP, which both reside predominantly in the C2'-endo conformation, have at least millimolar affinities.

Without the anchoring presence of Thr45, the inhibitor can orient its furanose ring and phosphoryl group so as to optimise favourable contacts with active-site residues (Fig 4.8a-b). Thus, 3'-UMP could form a hydrogen bond between its 2'-OH and His12 or other active-site residues instead of participating in sterically unfavourable interactions. This putative hydrogen bond could be the source of the two-fold higher affinity of T45G RNase A for 3'-UMP than dU^FMP. dU^FMP cannot form a hydrogen bond with its 2'-fluoro substituent, but has a favourable conformation and pK_a for binding to the active site at pH 6.

Although araUMP has the same phosphoryl group pK_a value as 3'-UMP and dU^{F}MP, it has at least an order of magnitude lower affinity for T45G RNase A. Ring pucker could explain the lack of affinity between araUMP and T45G RNase A (compared to 3'-UMP and dU^{F}MP). When the furanose ring is in the C2'-endo conformation, the nucleobase and phosphoryl group are placed in such a way that neither can interact optimally with active-site residues in the absence of the anchoring site provided by Thr45 in the wild-type enzyme. This lack of favourable interactions should be true for dUMP, as well. dUMP has the additional handicap of having a higher phosphoryl group pK_a than the other 3'-nucleotides, with consequently diminished Coulombic interactions.

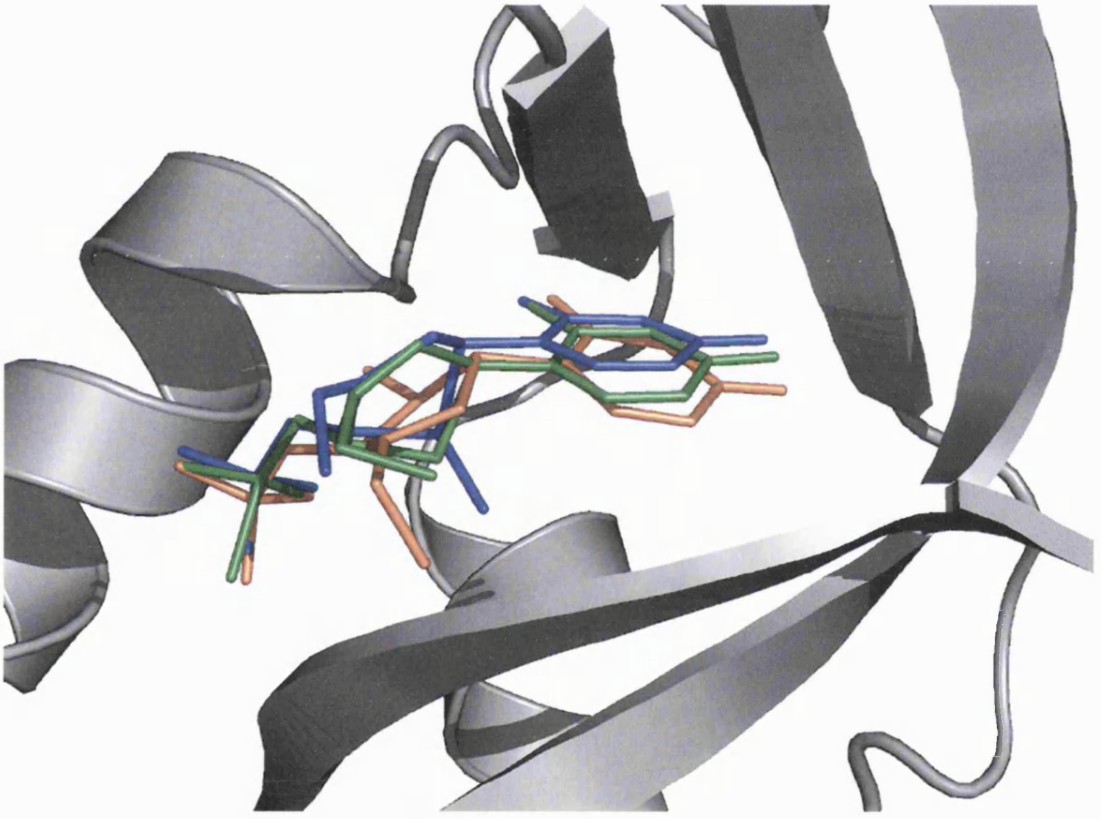


Fig 4.8a Anchoring of the phosphate moiety and base moiety at the active site leading to flexibility in ribose puckering.

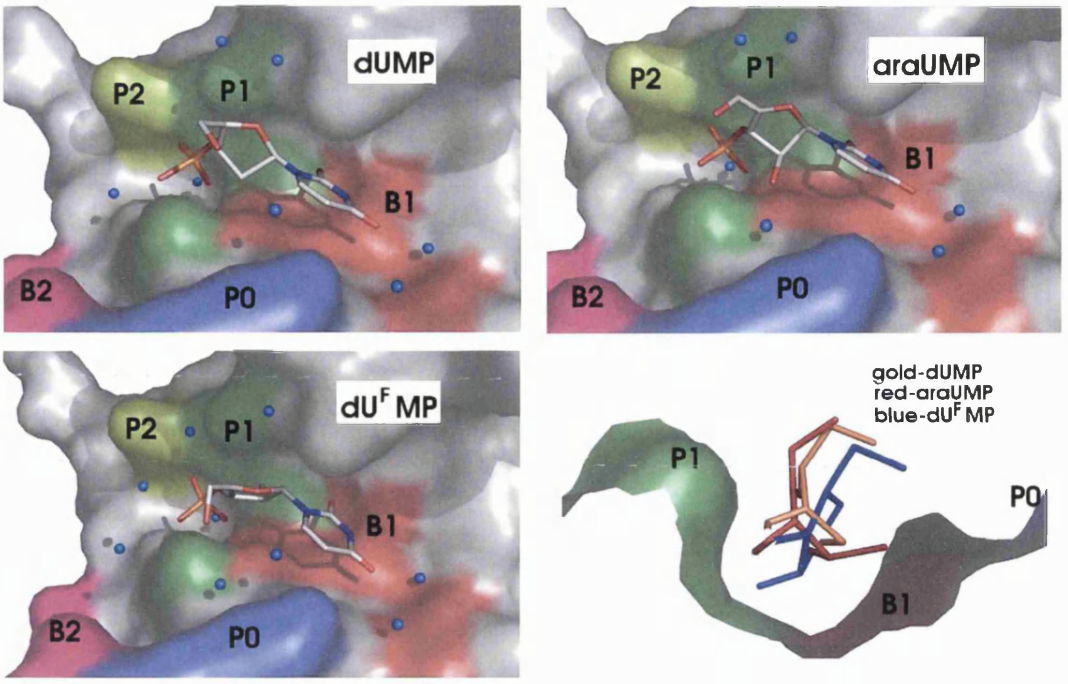


Fig 4.8b Surface representation of active sites of RNase A·non-natural 3'-nucleotide complex. All the subsites are shown. The puckering of ribose moiety is shown in the right bottom quadrant of the figure.

Conclusion

Though the furanose ring puckering is specific to the T45G variant, it is less sensitive to wild-type RNase A (Table 4.2). The two most effective inhibitors of wild-type RNase A reported in this work are dU^FMP and araUMP, which have different puckers in solution but indistinguishable K_i values. It has already been shown that in the context of a deoxyribose tetranucleotide ligand, a 2'-fluoro-2'-deoxyuridine residue binds more weakly to RNase A than does a 2'-deoxyuridine residue (Kelemen et al. 2000). Hence, we suggest that a most effective residue in creating new inhibitors of RNase A is arabinouridine, whose *pKa* value is close to dUMP and binds strongly to wild-type RNase A than dUMP. These compounds can be considered as the basis for the creation of new high-affinity ligands that could be used as effective drug molecules and aid in understanding the mechanism of RNase A homologues.

Appendix I

Short Definitions

Amino acid: *Amino acids are building blocks of proteins and there exist 20 different natural amino acids. Each amino acid contains an amino- and carboxyl- group attached to alpha carbon.*

Bioinformatics: *A new field in biology uses information technology to aid in solving biological problems! Some of its functions are data mining and statistics, and development of algorithms for use in gene finding, sequence alignment, protein structure prediction, and genome assembly and organisation.*

Bond angle: *For any three points A, B and C in space, where A is bonded to B and B to C, then the angle between AB and BC is called bond angle.*

Bond length: *For any two points A and B that are bonded, the length between them is called bond length.*

Chemical modification: *To alter the chemical property of the molecule by substitution, deletion or addition of atoms.*

Dihedral angle: *Angle between any two planes is called dihedral angle. Also called as torsion angle.*

Eukaryote: *Eukaryotes are multicellular organisms. Some examples are plants, animals, and insects.*

Genomics: A genome is a complete set of information about an organism at its DNA level. Genomics is the study of organisation of genomes to understand evolution, and characters of inheritance.

Microarray technology: *A new high-throughput molecular biology technique to map and characterise genes, single nucleotide polymorphism (SNP) of genes, etc,.of a genome in a chip.*

Nucleic acids: *Nucleic acids are building blocks of deoxy-ribonucleic acid (DNA) and ribonucleic acid (RNA), the genetic material of an organism. They are classified as purines and pyrimidines.*

Peptide bond: *The covalent bond between two amino acids is called peptide bond. The covalent bond is formed between carboxy- group of N-1th amino acid to amino- group of Nth amino acid. The geometry of the peptide bond is planar.*

Plasmid: *Plasmids are circular DNA material of prokaryotes and some eukaryotes independent of the chromosomal DNA of a cell. They are gene transfer vehicles between genomes. Genetically engineered plasmids are called as vectors and used extensively in research.*

Prokaryote: *Prokaryotes are single celled organisms. eg. bacteria.*

Protein: *Proteins are polymers of amino acids linked together by peptide bonds. They exist in different shape and structure and are synthesised in an organism. Proteins are vital for cellular functions.*

Proteomics: *The study of the proteome of an organism is called proteomics. High-throughput study involves structural characterisation of different proteins on a large scale using robotics, pattern analysis of proteins in diseases, etc.*

Virus: *Viruses are particles that are capable of infecting prokaryotic or eukaryotic host cells. They are host dependent and cannot replicate on their own, but contain genetic material.*

Appendix II

List of Publications

1. Lloyd, M. D., Thiyagarajan, N., Ho, Y. T.,Lawrence Woo, L. W., Sutcliffe, O. B., Purohit, A., Reed, N. J., Acharya, K. R. and Potter, B. V. L. (2005) First crystal structures of human carbonic anhydrase II in complex with dual aromatase-steroid sulfatase inhibitors. *Biochemistry*, **44**, 6858-6866.

2. Jenkins, C. L., Thiyagarajan, N., Sweeney, R. Y., Guy, M. P., Kelemen, B. R., Acharya, K. R. and Raines, R. T. (2005) Binding of non-natural 3'-nucleotides to ribonuclease A. *FEBS Journal*, **272**, 744-755.

3. Turton, K., Ramanathan, N., Thiyagarajan, N., Chaddock, J. A. and Acharya, K. R. (2004) Crystal structures of *Erythrina cristagalli* lectin with bound N-linked oligosaccharide and lactose. *Glycobiology*, **14**, 923-929.

References

Abel, S., and Glund, K. 1986. Localization of RNA-degrading enzyme activity within vacuoles of cultured tomato cells. *Physiol. Plant.* **66:** 79-86.

Acharya, K.R., and Rees, A.R. 1995. X-Ray Diffraction of Biomolecules. In: *Molecular Biology and Biotechnology,* Meyers, R.A.**:** 275-282.

Acharya, K.R., Shapiro, R., Allen, S.C., Riordan, J.F., and Vallee, B.L. 1994. Crystal structure of human angiogenin reveals the structural basis for its functional divergence from pancreatic ribonuclease. *Proc Natl Acad Sci U S A* **91:** 2915-2919.

Acharya, K.R., Shapiro, R., Riordan, J.F., and Vallee, B.L. 1995. Crystal structure of bovine angiogenin at 1.5Å resolution. *Proc Natl Acad Sci U S A* **92:** 2949-2953.

Acharya, K.R., Subramanian, V., Shapiro, R., Riordan, J.F., and Vallee, B.L. 1992. Crystallization and preliminary X-ray analysis of Human angiogenin. *J. Mol. Biol.* **228:** 1269-1270.

Agarwal, R., Binz, T., and Swaminathan, S. 2005. Analysis of Active Site Residues of Botulinum Neurotoxin E by Mutational, Functional, and Structural Studies: Glu335Gln Is an Apoenzyme. *Biochemistry* **44:** 8291-8302.

Ahmed, S.A., and Smith, L.A. 2000. Light chain of botulinum A neurotoxin expressed as an inclusion body from a synthetic gene is catalytically and functionally active. *J Protein Chem* **19:** 475-487.

Ahnert-Hilger, G., and Bigalke, H. 1995. Molecular aspects of tetanus and botulinum neurotoxin poisoning. *Prog Neurobiol* **46:** 83-96.

Ahnert-Hilger, G., Weller, U., Dauzenroth, M.E., Habermann, E., and Gratzl, M. 1989. The tetanus toxin light chain inhibits exocytosis. *FEBS Lett* **242:** 245-248.

Altona, C., and Sundaralingam, M. 1972. Conformational analysis of the sugar ring in nucleosides and nucleotides. A new description using the concept of pseudorotation. *J. Am. Chem. Soc.* **94:** 8205-8212.

Anne, C., Blommaert, A., Turcaud, S., Martin, A.-S., Meudal, H., and Roques, B.P. 2003a. Thio-derived disulfides as potent inhibitors of botulinum

neurotoxin type B: Implications fir zinc interaction. *Bioorg Med Chem* **11:** 4655-4660.

Anne, C., Turcaud, S., Quancard, J., Teffo, F., Meudal, H., Fournie-Zaluski, M.-C., and Roques, B.P. 2003b. Development of potent inhibitors of botulinum neurotoxin type B. *J. Med. Chem.* **46:** 4648-4656.

Annese, V., Basciani, M., Borrelli, O., Leandro, G., Simone, P., and Andriulli, A. 1998. Intrasphincteric injection of botulinum toxin is effective in long-term treatment of esophageal achalasia. *Muscle Nerve* **21:** 1540-1542.

Antonov, I.V., Dudkin, S.M., Karpeiskii, M.Y., and Yakovlev, G.I. 1976. The conformations of phosphorylating derivatives of 2′-fluoro-2′-deoxyuridine in solution. *Sov. J. Bioorg. Chem.* **2:** 863-872.

Ardelt, W., Mikulski, M., and Shogen, K. 1991. Amino acid sequence of an anti-tumor protein from *Rana pipiens* oocytes and early embryos. *J. Biol. Chem.* **266:** 245-251.

Ashford, D., Dwek, R.A., Welply, J.K., Amatayakul, S., Homans, S.W., Lis, H., Taylor, G.N., Sharon, N., and Rademacher, T.W. 1987. The beta 1----2-D-xylose and alpha 1----3-L-fucose substituted N-linked oligosaccharides from Erythrina cristagalli lectin. Isolation, characterisation and comparison with other legume lectins. *Eur J Biochem* **166:** 311-320.

Ashford, D.A., Dwek, R.A., Rademacher, T.W., Lis, H., and Sharon, N. 1991. The glycosylation of glycoprotein lectins. Intra- and inter-genus variation in N-linked oligosaccharide expression. *Carbohydr Res* **213:** 215-227.

Aslanidis, C., and de Jong, P.J. 1990. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res* **18:** 6069-6074.

Bajjalieh, S. 2001. SNAREs take the stage: a prime time to trigger neurotransmitter secretion. *TRENDS in Neurosciences* **24:** 678-680.

Barnard, E.A., and Stein, W.D. 1959. The histidine residue in the active center of ribonuclease. I. A specific reaction with bromoacetic acid. *J Mol Biol* **1:** 339-349.

Beers, W.H., and Reich, E. 1968. Isolation and charaterization of *Clostridium botulinum* type B toxin. *J. Biol. Chem.* **244:** 4473-4479.

Beintema, J.J. 1998. Introduction: the ribonuclease A superfamily. *Cell Mol Life Sci* **54:** 763-765.

Beintema, J.J., and Neuteboom, B. 1983. Origin of the duplicated ribonuclease gene in guinea-pig: comparison of the amino acid sequences with those of two close relatives: capybara and cuis ribonuclease. *J Mol Evol* **19:** 145-152.

Beintema, J.J., Schuller, C., Irie, M., and Carsana, A. 1988. Molecular evolution of the ribonuclease superfamily. *Prog Biophys Mol Biol* **51:** 165-192.

Bicknell, R., and Vallee, B.L. 1988. Angiogenin activates endothelial cell phospholipase C. *Proc Natl Acad Sci U S A* **85:** 5961-5965.

Bicknell, R., and Vallee, B.L. 1989. Angiogenin stimulates endothelial cell prostacyclin secretion by activation of phospholipase A2. *Proc Natl Acad Sci U S A* **86:** 1573-1577.

Binz, T., Blasi, J., Yamasaki, S., Baumeister, A., Link, E., Sudhof, T.C., Jahn, R., and Niemann, H. 1994. Proteolysis of SNAP-25 by Types E and A botulinum neurotoxins. *J. Biol. Chem.* **269:** 1617-1620.

Binz, T., Kurazono, H., Wille, M., Frevert, J., Wernars, K., and Niemann, H. 1990. The complete sequence of botulinum neurotoxin type A and comparison with other clostridial neurotoxins. *J Biol Chem* **265:** 9153-9158.

Birdsall, D.L., and McPherson, A. 1992. Crystal structure disposition of thymidylic acid tetramer in complex with ribonuclease A. *J. Biol. Chem.* **267:** 22230-22236.

Black, J.D., and Dolly, J.O. 1986a. Interaction of 125I-labeled botulinum neurotoxins with nerve terminals. I. Ultrastructural autoradiographic localization and quantitation of distinct membrane acceptors for types A and B on motor nerves. *J Cell Biol* **103:** 521-534.

Black, J.D., and Dolly, J.O. 1986b. Interaction of 125I-labeled botulinum neurotoxins with nerve terminals. II. Autoradiographic evidence for its uptake into motor nerves by acceptor-mediated endocytosis. *J Cell Biol* **103:** 535-544.

Blaustein, R.O., Germann, W.J., Finkelstein, A., and DasGupta, B.R. 1987. The N-terminal half of the heavy chain of botulinum type A neurotoxin forms channels in planar phospholipid bilayers. *FEBS Lett* **226:** 115-120.

Blow, D.M., and Rossman, M.G. 1961. The Single Isomorphous Replacement Method. *Acta Crystallogr* **14:** 1195-1203.

Blundell, T.L., and Johnson, L.N. 1976. In Protein Crystallography. *Academic Press, London.*

Boix, E., Leonidas, D.D., Nikolovski, Z., Nogues, M.V., Cuchillo, C.M., and Acharya, K.R. 1999. Crystal structure of eosinophil cationic protein at 2.4 Å resolution. *Biochemistry* **38:** 16794-16801.

Boix, E., Wu, Y., Vasandani, V.M., Saxena, S.K., Ardelt, W., Ladner, J., and Youle, R.J. 1996. Role of the N terminus in RNase A homologues: differences in catalytic activity, ribonuclease inhibitor interaction and cytotoxicity. *J Mol Biol* **257:** 992-1007.

Bond, M.D., Strydom, D.J., and Vallee, B.L. 1993. Characterization and sequencing of rabbit, pig and mouse angiogenins: discernment of functionally important residues and regions. *Biochim Biophys Acta* **1162:** 177-186.

Boyd, W.C., and Reguera, R.M. 1949. Studies on haemagglutinins present in seeds of some representatives of the family Leguminoseae. *J. Immunol.* **62:** 333-339.

Bragg, W.H. 1913a. The reflection of X-rays by crystals. *Proc. Roy. Soc. London (A)* **88:** 428-438.

Bragg, W.L. 1913b. The diffraction of short electromagnetic waves by a crystal. *Proc. Cambridge. Phil. Soc.* **17:** 43-57.

Breidenbach, M.A., and Brunger, A.T. 2004. Substrate recognition strategy for botulinum neurotoxin serotype A. *Nature* **432:** 925-929.

Brunger, A.T. 1992. Free R value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355:** 472-475.

Brunger, A.T. 1993. Assessment of phase accuracy by cross calidation: the free *R* value. Methods and applications. *Acta Crystallogr* **D49:** 24-36.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D* **54:** 905-921.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D* **54:** 905-921.

Burgen, A.S.V., Dickens, F., and Zatman, L.J. 1949. The action of botulinum toxin on the neuromuscular junction. *J Physiol (lond)* **109:** 10-24.

Burnett, J.C., Schmidt, J.J., Stafford, R.G., Panchal, R.G., Nguyen, T.L., Hermone, A.R., Vennerstrom, J.L., McGrath, C.F., Lane, D.J., Sausville, E.A., et al. 2003. Novel small molecule inhibitors of botulinum neurotoxin A metalloprotease activity. *Biochem Biophys Res Commun* **310:** 84-93.

Chaddock, J.A., Herbert, M.H., Ling, R.J., Alexander, F.C.G., Fooks, S.J., Revell, D.F., Quinn, C.P., Shone, C.C., and Foster, K.A. 2002. Expression and purification of catalytically active, non-toxic endopeptidase derivatives of *Clostridium botulinum* toxin type A. *Protein Expr Purif* **25:** 219-228.

Chaddock, J.A., Purkiss, J.R., Alexander, F.C., Doward, S., Fooks, S.J., Friis, L.M., Hall, Y.H., Kirby, E.R., Leeds, N., Moulsdale, H.J., et al. 2004. Retargeted clostridial endopeptidases: inhibition of nociceptive neurotransmitter release in vitro, and antinociceptive activity in in vivo models of pain. *Mov Disord* **19:** S42-S47.

Chaddock, J.A., Purkiss, J.R., Duggan, M.J., Quinn, C.P., Shone, C.C., and Foster, K.A. 2000a. A conjugate composed of nerve growth factor coupled to a non-toxic derivative of *Clostridium botulinum* neurotoxin type A can inhibit neurotransmitter release. *Growth Factors* **18:** 147-155.

Chaddock, J.A., Purkiss, J.R., Friis, L.M., Broadbridge, J.D., Duggan, M.J., Fooks, S.J., Shone, C.C., Quinn, C.P., and Foster, K.A. 2000b. Inhibition of vesicular secretion in both neuronal and non-neuronal cells by a retargeted ednopeptidase dervative of *Clostridium botulinum* neurotoxin type A. *Infect Immun* **68:** 2587-2593.

Chen, Y.A.*e.a.* 2001. Sequential SNARE assembly underlies priming and triggering of exocytosis. *Neuron* **30:** 161-170.

Cho, S., Beintema, J.J., and Zhang, J. 2005. The ribonuclease A superfamily of mammals and birds: identifying new members and tracing evolutionary histories. *Genomics* **85:** 208-220.

Chong, S., Mersha, F.B., Comb, D.G., Scott, M.E., Landry, D., Vence, L.M., Perler, F.B., Benner, J., Kucera, R.B., Hirvonen, C.A., et al. 1997. Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* **192:** 271-281.

Chong, S., Montello, G.E., Zhang, A., Cantor, E.J., Liao, W., Xu, M.Q., and Benner, J. 1998a. Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. *Nucleic Acids Res* **26:** 5109-5115.

Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F.B., and Xu, M.Q. 1996. Protein splicing involving the Saccharomyces cerevisiae VMA intein. The steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an in vitro splicing system. *J Biol Chem* **271:** 22159-22168.

Chong, S., Williams, K.S., Wotkowicz, C., and Xu, M.Q. 1998b. Modulation of protein splicing of the Saccharomyces cerevisiae vacuolar membrane ATPase intein. *J Biol Chem* **273:** 10567-10577.

Chwang, A.K., and Sundaralingam, M. 1973. Intramolecular hydrogen bonding in 1-β-D-arabinofuranosylcytosine (Ara-C). *Nat. New Biol.* **243:** 78-79.

Colman, P.M., Jansonius, J.N., and Matthews, B.W. 1972. The structure of thermolysin: an electron density map at 2-3 A resolution. *J Mol Biol* **70:** 701-724.

Cronan, J.E., Jr. 1990. Biotination of proteins in vivo. A post-translational modification to label, purify, and study proteins. *J Biol Chem* **265:** 10327-10333.

Cuchillo, C.M., Moussaoui, M., Barman, T., Travers, F., and Nogues, M.V. 2002. The exo- or endonucleolytic preference of bovine pancreatic ribonuclease A depends on its subsites' structure and on the substrate size. *Protein Sci.* **11:** 117-128.

D'Alessio, G., and Riordan, G.F. 1997. *Ribonucleases: Structures and Functions.* Academic Press, New York.

Dasgupta, B.R., and Sugiyama, H. 1972a. A common subunit structure in Clostridium botulinum type A, B and E toxins. *Biochem Biophys Res Commun* **48:** 108-112.

Dasgupta, B.R., and Sugiyama, H. 1972b. Isolation and characterization of a protease from Clostridium botulinum type B. *Biochim Biophys Acta* **268:** 719-729.

Davies, D.B., and Danyluk, S.S. 1975. Nuclear magnetic resonance studies of 2′- and 3′-ribonucleotide structures in solution. *Biochemistry* **14:** 543-554.

Davis, I.W., Murray, L.W., Richardson, J.S., and Richardson, D.C. 2004. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Research.* **32:** W615-W619.

De Boeck, H., Macgregor, R.B., Jr., Clegg, R.M., Sharon, N., and Loontiens, F.G. 1985. Binding of N-dansylgalactosamine to the lectin from Erythrina cristagalli as followed by stopped-flow and pressure-jump relaxation kinetics. *Eur J Biochem* **149:** 141-145.

delCardayré, S.B., and Raines, R.T. 1994. Structural determinants of enzymatic processivity. *Biochemistry* **33:** 6031-6037.

delCardayré, S.B., and Raines, R.T. 1995. A residue to residue hydrogen bond mediates the nucleotide specificity of ribonuclease A. *J. Mol. Biol.* **252:** 328-336.

Derewenda, Z., Yariv, J., Helliwell, J.R., Kalb, A.J., Dodson, E.J., Papiz, M.Z., Wan, T., and Campbell, J. 1989. The structure of the saccharide-binding site of concanavalin A. *Embo J* **8:** 2189-2193.

Deshpande, S.S., Sheridan, R.E., and Adler, M. 1997. Efficacy of certain quinolines as pharmacological antagonists in botulinum neurotoxin poisoning. *Toxicon* **35:** 433-445.

di Guan, C., Li, P., Riggs, P.D., and Inouye, H. 1988. Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein. *Gene* **67:** 21-30.

Donovan, J.J., and Middlebrook, J.L. 1986. Ion-conducting channels produced by botulinum toxin in planar lipid membranes. *Biochemistry* **25:** 2872-2876.

Duggan, M.J., Quinn, C.P., Chaddock, J.A., Purkiss, J.R., Alexander, F.C.G., Doward, S., Fooks, S.J., Friis, L.M., Hall, Y.H.J., Kirby, E.R., et al. 2002. Inhibition of Release of Neurotransmitters from Rat Dorsal Root Ganglia by a Novel Conjugate of a Clostridium botulinum Toxin A Endopeptidase Fragment and Erythrina cristagalli Lectin. *J Biol Chem* **277:** 34846-34852.

Durack, D.T., Ackerman, S.J., Loegering, D.A., and Gleich, G.J. 1981. Purification of human eosinophil-derived neurotoxin. *Proc Natl Acad Sci U S A* **76:** 1443-1447.

Dyer, K.D., and Rosenberg, H.F. 2006. The RNase a superfamily: Generation of diversity and innate host defense. *Mol Divers.*

Elgavish, S., and Shaanan, B. 1998. Structures of the *Erythrina corallodendron* lectin and of its complexes with mono- and disaccharides. *J. Mol. Biol.* **277:** 917-932.

Elgavish, S., and Shaanan, B. 2001. Chemical characteristics of dimer interfaces in the legume lectin family. *Prot. Sci.* **10:** 753-761.

Emmerich, C., Helliwell, J.R., Redshaw, M., Naismith, J.H., Harrop, S.J., Raftery, J., Kalb, A.J., Yariv, J., Dauter, Z., and Wilson, K.S. 1994. High-resolution structures of single-metal-substituted concanavalin A: the Co, Ca-protein at 1.6Å and the Ni, Ca-protein at 2.0Å. *Acta Crystallogr.* **D50:** 749-756.

Esnouf, R.M. 1997a. An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J Mol Graph Model* **15:** 132-134.

Esnouf, R.M. 1997b. An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J Mol Graph Model* **15(2):** 132-134.

Eswaramoorthy, S., Kumaran, D., Keller, J., and Swaminathan, S. 2004. Role of metals in the biological activity of Clostridium botulinum neurotoxins. *Biochemistry* **43:** 2209-2216.

Fisher, B.M., Grilley, J.E., and Raines, R.T. 1998a. A new remote subsite in ribonuclease A. *J. Biol. Chem.* **273:** 34134-34138.

Fisher, B.M., Ha, J.-H., and Raines, R.T. 1998b. Coulombic forces in protein-RNA interactions: Binding and cleavage by ribonuclease A and variants at Lys7, Arg10 and Lys66. *Biochemistry* **37:** 12121-12132.

Fleer, R., Yeh, P., Amellal, N., Maury, I., Fournier, A., Bacchetta, F., Baduel, P., Jung, G., L'Hote, H., Becquart, J., et al. 1991. Stable multicopy vectors for high-level secretion of recombinant human serum albumin by Kluyveromyces yeasts. *Biotechnology (N Y)* **9:** 968-975.

Fontecilla-Camps, J.C., de Llorens, R., le Du, M.H., and Cuchillo, C.M. 1994. Crystal structure of ribonuclease A.d(ApTpApApG) complex. *J. Biol. Chem.* **269:** 21526-21531.

Foster, K.A., Adams, E.J., Durose, L., Cruttwell, C.J., Marks, E., Shone, C.C., Chaddock, J.A., Cox, C.L., Heaton, C., Sutton, J.M., et al. 2006. Re-engineering the target specificity of clostridial neurotoxins - a route to novel therapeutics. *Neurotoxicity Res.* **9:** 101-107.

Giacovazzo, C., Siliqi, D., and García-Rodríguez, L. 2001. On integrating direct methods and isomorphous-replacement techniques: triplet estimation and treatment of errors. *Acta Crystallogr* **A57:** 571-575.

Ginalski, K., Venclovas, C., Lesyng, B., and Fidelis, K. 2000. Structure-based sequence alignment for the beta-trefoil subdomain of the clostridial neurotoxin family provides residue level information about the putative ganglioside binding site. *FEBS Lett* **482:** 119-124.

Gleich, G.J., Loegering, D.A., Bell, M.P., Checkel, J.L., Ackerman, S.J., and McKean, D.J. 1986. Biochemical and functional similarities between human eosinophil-derived neurotoxin and eosinophil cationic protein: homology with ribonuclease. *Proc Natl Acad Sci U S A* **83:** 3146-3150.

Gomis-Ruth, F.X., Kress, L.F., Kellermann, J., Mayr, I., Lee, X., Huber, R., and Bode, W. 1994. Refined 2.0 A X-ray crystal structure of the snake venom zinc-endopeptidase adamalysin II. Primary and tertiary structure determination, refinement, molecular structure and comparison with astacin, collagenase and thermolysin. *J Mol Biol* **239:** 513-544.

Graham, F.L., and van der Eb, A.J. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52:** 456-467.

Green, D.W., Ingram, V.M., and Perutz, M.F. 1954. The structure of haemoglobin IV. Sign determination by the isomorphous replacement method. *Proc. Roy. Soc. London (A)* **225:** 287-307.

Greene, P., Kang, U., Fahn, S., Brin, M., Moskowits, C., and Flaster, E. 1990. Double-blind, placebo-controlled trial of botulinum toxin injections for the treatmetn of spasmodic torticollis. *Neurology* **40:** 1213-1218.

Gundlach, H.G., Stein, W.H., and Moore, S. 1959. The nature of the amino acid residues involved in the inactivation of ribonuclease by iodoacetate. *J Biol Chem* **234:** 1754-1760.

Guschlbauer, W., and Jankowski, K. 1980. Nucleoside conformation is determined by the electronegativity of the sugar substituent. *Nucleic Acids Res.* **8:** 1421-1433.

Hamann, K.J., Gleich, G.J., Checkel, J.L., Loegering, D.A., McCall, J.W., and Barker, R.L. 1990. In vitro killing of microfilariae of Brugia pahangi and Brugia malayi by eosinophil granule proteins. *J Immunol* **144:** 3166-3173.

Hanhn, T., Vaos, A., and Buerger, M.J. 1987. *International Tables For Crystallography*, Second, revised edition ed. D. Reidel Publishing Company, pp. 11-67.

Harris, D.T., Iglesias, J.L., Argov, S., Toomey, J., and Koren, H.S. 1987. Heterogeneity of human natural killer (NK) cells: enrichment of NK by

negative-selection with the lectin from Erythrina cristagalli. *J Leukoc Biol* **42:** 163-170.

Harrison, A.M., Bonville, C.A., Rosenberg, H.F., and Domachowske, J.B. 1999. Respiratory syncytical virus-induced chemokine expression in the lower airways: eosinophil recruitment and degranulation. *Am J Respir Crit Care Med* **159:** 1918-1924.

Haynes, L.P., Morgan, A., and Burgoyne, R.D. 1999. nSec-1 (munc-18) interacts with both primed and unprimed syntaxin 1A and associates in a dimeric complex on adrenal chromaffin granules. *Biochem J* **342 Pt 3:** 707-714.

Hendrickson, W.A., Smith, J.L., and Sheriff, S. 1985. Direct phase determination based on anomalous scattering. *Methods Enzymol.* **115:** 41-55.

Hofsteenge, J., Vicentini, A., and Zelenko, O. 1998. Ribonuclease 4, an evolutionarily highly conserved member of the superfamily. *Cell Mol Life Sci* **54:** 804-810.

Holden, H.M., Tronrud, D.E., Monzingo, A.F., Weaver, L.H., and Matthews, B.W. 1987. Slow- and fast- binding inhibitors of thermolysin display different modes of binding: crystallographic analysis of extended phophonamidate transition-state analogues. *Biochemistry* **26:** 8542-8553.

Holloway, D.E., Chavali, G.B., Hares, M.C., Baker, M.D., Subbarao, G.V., Shapiro, R., and Acharya, K.R. 2004. Crystallographic studies on structural features that determine the enzymatic specificity and potency of human angiogenin: Thr44, Thr80, and residues 38-41. *Biochemistry* **43:** 1230-1241.

Holloway, D.E., Chavali, G.B., Hares, M.C., Subramanian, V., and Acharya, K.R. 2005. Structure of murine angiogenin: features of the substrate- and cell-binding regions and prospects for inhibitor-binding studies. *Acta Crystallogr* **D61:** 1568-1578.

Holloway, D.E., Hares, M.C., Shapiro, R., Subramanian, V., and Acharya, K.R. 2001. High-level expression of three members of the murine

Angiogenin family in Escherichia coli and purification of the recombinant proteins. *Prot. Expr. Purif.* **22:** 307-317.

Howard, J.A.K., Hoy, V.J., O'Hagan, D., and Smith, G.T. 1996. How good is fluorine as a hydrogen bond acceptor? *Tetrahedron* **52:** 12613-12622.

Hu, G., Xu, C., and Riordan, J.F. 2000. Human angiogenin is rapidly translocated to the nucleus of human umbilical vein endothelial cells and binds to DNA. *J Cell Biochem* **76:** 452-462.

Huhn, M., Sasse, S., Tur, M.K., Matthey, B., Schinkothe, T., Rybak, S.M., Barth, S., and Engert, A. 2001. Human angiogenin fused to human CD30 ligand (Ang-CD30L) exhibits specific cytotoxicity against CD30-positive lymphoma. *Cancer Res* **61:** 8737-8742.

Humeau, Y. 2000. How botulinum and tetanus neurotoxins block neurotransmitter release. *Biochimie* **82:** 427-446.

Iglesias, J.L., Lis, H., and Sharon, N. 1982. Purification and properties of a D-galactose/N-acetyl-D-galactosamin-specific lectin from *Erythrina cristagalli. Eur. J. Biochem.* **123:** 247-252.

Iwasaki, M., and Sakaguchi, G. 1978. Acid preparation of *Clostridium botulinum* type C and D toxin from whole culture by addition of ribonucleic acid as precipitation aid. *Infect Immun* **19:** 749-751.

Iyer, S., Holloway, D.E., Kumar, K., Shapiro, R., and Acharya, K.R. 2005. Molecular recognition of human eosinophil derived neurotoxin (RNase-2) by placental ribonuclease inhibitor. *J Mol Biol* **347:** 637-655.

Jankovic, J., and Orman, J. 1987. Botulinum A toxin for cranial-cervical dystonia: a double-blind, placebo-controlled study. *Neurology* **37:** 616-623.

Jardine, A.M., Leonidas, D.D., Jenkins, J.L., Park, C., Raines, R.T., Acharya, K.R., and Shapiro, R. 2001. Cleavage of 3',5'-pyrophosphate-linked dinucleotides by ribonuclease A and angiogenin. *Biochemistry* **40:** 10262-10272.

Jenkins, C.L., Thiyagarajan, N., Sweeney, R.Y., Guy, M.P., Kelemen, B.R., Acharya, K.R., and Raines, T. 2005. Binding of non-natural 3'-nucleotides to ribonuclease A. *FEBS Journal* **272:** 744-755.

Jimi, S., Ito, K., Kohno, K., Ono, M., Kuwano, M., Itagaki, Y., and Ishikawa, H. 1995. Modulation by bovine angiogenin of tubular morphogenesis and expression of plasminogen activator in bovine endothelial cells. *Biochem Biophys Res Commun* **211:** 476-483.

Jinno, H., Ueda, M., Ozawa, S., Ikeda, T., Enomoto, K., Psarras, K., Kitajima, M., Yamada, H., and Seno, M. 1996. Epidermal growth factor receptor-dependent cytotoxicity for human squamous carcinoma cell lines of a conjugate composed of human EGF and RNase 1. *Life Sci* **58:** 1901-1908.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. 1991. Improved methods for building models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47:** 110-119.

Kagan, B.L., Finkelstein, A., and Colombini, M. 1981. Diptheria toxin fragment forms large pores in phospholipids bilayer membranes. *Proc Natl Acad Sci U S A* **78:** 4950-4954.

Kalandakanond, S., and Coffield, J.A. 2001. Cleavage of SNAP-25 by botulinum toxin type A requires receptor-mediated endocytosis, pH-dependent translocation and zinc. *J Pharmacol Exp Ther* **296:** 980-986.

Kao, I., Drachman, D.B., and Price, D.L. 1976. Botulinum toxin: Mechanism of presynaptic blockade. *Science (Wash DC)* **192:** 1256-1258.

Katayanagi, K., Okumura, M., and Morikawa, K. 1993. Crystal structure of Escherichia coli RNase HI in complex with Mg2+ at 2.8 Å resolution: proof for a single Mg(2+)-binding site. *Proteins* **17:** 337-346.

Kelemen, B.R., Klink, T.A., Behlke, M.A., Eubanks, S.R., Leland, P.A., and Raines, R.T. 1999. Hypersensitive substrate for ribonucleases. *Nucleic Acids Res.* **27:** 3696-3701.

Kelemen, B.R., Schultz, L.W., Sweeney, R.Y., and Raines, R.T. 2000. Excavating an active site: The nucleobase specificity of ribonuclease A. *Biochemistry* **39:** 14487-14494.

Kitamura, M., Sakaguchi, S., and Sakaguchi, G. 1968. Purification and some properties of *Clostridium botulinum* type-E toxin. *Biochim. Biophys. Acta* **168:** 207-217.

Kleywegt, G., and Jones, T.A. 1998a. Databases in protein crystallography. *Acta Crystallogr* **D54:** 1119-1131.

Kleywegt, G.J., and Jones, T.A. 1998b. Databases in protein crystallography. *Acta Crystallogr D* **54:** 1119-1131.

Klion, A.D., and Nutman, T.B. 2004. The role of eosinophils in host defense against helminth parasites. *J. Allergy Clin. Immunol.* **113:** 30-37.

Kobe, B., and Deisenhofer, J. 1993. Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature* **366:** 751-756.

Kobe, B., and Deisenhofer, J. 1995. A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature* **374:** 183-186.

Kobe, B., and Deisenhofer, J. 1996. Mechanism of ribonuclease inhibition by ribonuclease inhibitor protein based on the crystal structure of its complex with ribonuclease A. *J Mol Biol* **264:** 1028-1043.

Konieczny, A., Voytas, D.F., Cummings, M.P., and Ausubel, F.M. 1991. A superfamily of Arabidopsis thaliana retrotransposons. *Genetics* **127:** 801-809.

Kozaki, S., Sakaguchi, S., and Sakaguchi, G. 1974. Purification and some properties of proenitor toxins of *Clostridium botulinum* type B. *Infect Immun* **10:** 750-756.

Kunitz, M. 1940. Crystalline Ribonuclease. *J. gen. Physiol.* **24:** 15-32.

Lacy, D.B., and Stevens, R.C. 1999. Sequence homology and structural analysis of the clostridial neurotoxins. *J Mol Biol* **291:** 1091-1104.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680-685.

Lalli, G., Bohnert, S., Deinhardt, K., Verastegui, C., and Schiavo, G. 2003. The journey of tetanus and botulinum neurotoxins in neurons. *Trends Microbiol* **11:** 431-437.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26:** 283-291.

Lee, J.J., and Lee, N.A. 2005. Eosinophil degranulation: an evolutionary vestige or a universally destructive effector function? *Clin. Exp. Allergy* **35:** 986-994.

Leggiadro, R.J. 2000. The threat of biological terrorism: a public health and infection control reality. *Infect. Control Hosp. Epidemiol.* **21:** 53-56.

Lehrer, R.I., Szklarek, D., Barton, A., Ganz, T., Hamann, K.J., and Gleich, G.J. 1989. Antibacterial properties of eosinophil major basic protein and eosinophil cationic protein. *J Immunol* **142:** 4428-4434.

Leland, P.A., Schultz, L.W., Kim, B.M., and Raines, R.T. 1998. Ribonuclease A variants with potent cytotoxic activity. *Proc Natl Acad Sci U S A* **95:** 10407-10412.

Leonidas, D.D., Boix, E., Prill, R., Suzuki, M., Turton, R., Minson, K., Swaminathan, G.J., Youle, R.J., and Acharya, K.R. 2001a. Mapping the ribonucleolytic active-site of eosinophil-derived neurotoxin (EDN): High resolution crystal structures of EDN complexes with adenylic nucleotide inhibitors. *J. Biol. Chem.* **276:** 15009-15017.

Leonidas, D.D., Chavali, G.B., Jardine, A.M., Li, S., Shapiro, R., and Acharya, K.R. 2001b. Binding of phosphate and pyrophosphate ions at the active site of human angiogenin as revealed by X-ray crystallography. *Protein Sci* **10:** 1669-1676.

Leonidas, D.D., Chavali, G.B., Oikonomakos, N.G., Chrysina, E.D., Kosmopoulou, M.N., Vlassi, M., Frankling, C., and Acharya, K.R. 2003. High resolution crystal structures of ribonuclease A complexed with adenylic and uridylic nucleotide inhibitors. Implications for structure-based design of ribonucleolytic inhibitors. *Protein Sci.* **12:** 2559-2574.

Leonidas, D.D., Shapiro, R., Allen, S.C., Subbarao, G.V., Veluraja, K., and Acharya, K.R. 1999a. Refined crystal structures of native human angiogenin and two active site variants: Implications for the unique functional properties of an enzyme involved in neovascularisation during tumour growth. *J Mol Biol* **285:** 1209-1233.

Leonidas, D.D., Shapiro, R., Irons, L.I., Russo, N., and Acharya, K.R. 1997. Crystal structures of ribonuclease A complexes with 5'-diphosphoadenosine 3'-phosphate and 5'-diphosphoadenosine 2'-phosphate at 1.7 Å resolution. *Biochemistry* **36:** 5578-5588.

Leonidas, D.D., Shapiro, R., Irons, L.I., Russo, N., and Acharya, K.R. 1999b. Toward rational design of ribonuclease inhibitors: High resolution

crystal structure of a ribonuclease A complex with a potent 3′,5′-pyrophosphate-linked dinucleotide inhibitor. *Biochemistry* **38:** 10287-10297.

Lowenthal, J.P., and Lamanna, C. 1953. Characterization of botulinal hemagglutination. *Am. J. Hyg.* **57:** 46-59.

Maina, C.V., Riggs, P.D., Grandea, A.G., 3rd, Slatko, B.E., Moran, L.S., Tagliamonte, J.A., McReynolds, L.A., and Guan, C.D. 1988. An Escherichia coli vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein. *Gene* **74:** 365-373.

Maksymowych, A.B., and Simpson, L.L. 1998. Binding and transcytosis of botulinum neurotoxin by polarized human colon carcinoma cells. *J Biol Chem* **273:** 21950-21957.

Manoj, N., and Suguna, K. 2001. Signature of quarternary structure in the sequences of legume lectins. *Protein Eng.* **14:** 735-745.

Maria, G., Cassetta, E., Gui, D., Brisinda, G., Bentivoglio, A.R., and Albanese, A. 1998. A comparison of botulinum toxin and saline for the treatment of chronic anal fissure. *N Engl J Med.* **338:** 217-220.

Matthews, B.W., Weaver, L.H., and Kester, W.R. 1974. The conformation of thermolysin. *J Biol Chem* **249:** 8030-8044.

McPherson, A., Brayer, G., Cascio, D., and Williams, R. 1986. The mechanism of binding of a polynucleotide chain to pancreatic ribonuclease. *Science* **232:** 765-768.

Mikhailov, V.V., and Mikhailov, V.V. 1970. Role of halogens in the molecule of some neuroleptics which prolong the survival of mice poisoned by botulinum toxin A. *Arch. Hig. Rada. Toksikol.* **55:** 143-145.

Minton, N.P., Machline, M.L., Lemmon, M.J., Brehm, J.K., Fox, M., Michael, N.P., Giaccia, A., and Brown, J.M. 1995. Chemotherapeutic tumour targeting using clostridial spores. *FEMS Microbiol Reviews* **17:** 357-364.

Miyazaki, S., Iwasaki, M., and Sakaguchi, G. 1977. *Clostridium botulinum* type D toxin: purification, molecular structure, and some immunological properties. *Infect Immun* **17:** 395-401.

Montecucco, C. 1986. How do tetanus and botulinum toxins bindi to neuronal membranes? *Trends Biochem Sci* **11:** 314-317.

Montecucco, C. 1994. Bacterial protein toxins penetrate cells via a four-step mechanism. *FEBS Lett.* **346:** 92-98.

Montecucco, C., and Schiavo, G. 1993. Tetanus and botulism neurotoxins: a new group of zinc proteases. *Trends Biochem Sci* **18:** 324-327.

Moreno, E., Teneberg, S., Adar, R., Sharon, N., Karlson, K.A., and Angstrom, J. 1997. Redefinition of the carbohydrate specificity of *Erythrina corallodendron* lectin based on solid-phase binding assays and molecular modeling of native and recombinant forms obtained by site-directed mutagenesis. *Biochemistry* **36:** 4429-4437.

Naumann, M., Jost, W.H., and Toyka, K.V. 1999. Botulinum toxin in the treatment of neurological disorders of the autonomic nervous system. *Arch Neurol* **56:** 914-916.

Navaza, J. 1994. AMoRe: an automated package for molecular replacement. *Acta Crystallogr* **A50:** 157-163.

Newton, D.L., Hansen, H.J., Mikulski, S.M., Goldenberg, D.M., and Rybak, S.M. 2001. Potent and specific antitumor effects of an anti-CD22-targeted cytotoxic ribonuclease: potential for the treatment of non-Hodgkin lymphoma. *Blood* **97:** 528-535.

Nogues, M.V., Moussaoui, M., Biox, E., Vilanova, M., Ribo, M., and Cuchillo, C.M. 1998. The contribution of noncatalytic phosphate-binding subsites to the mechanism of bovine pancreatic ribonuclease A. *Cell. Mol. Life Sci.* **54:** 766-774.

Oda, Y., Nakamura, H., Kanaya, S., and Ikehara, M. 1991. Binding of metal ions to E. coli RNase HI observed by 1H-15N heteronuclear 2D NMR. *J Biomol NMR* **1:** 247-255.

Ohishi, I., and Sakaguchi, G. 1975. Molecular construction of *Clostridium botulinum* type F progenitor toxin. *Appl. Microbiol.* **29:** 444-447.

Oppenheim, J.J., and Yang, D. 2005. Alarmins: chemotactic activators of immune responses. *Curr. Opin. Immunol.* **17:** 359-365.

Otwinowski, Z., and Minor, W. 1997a. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276:** 307-326.

Otwinowski, Z., and Minor, W. 1997b. Processing of X-ray diffraction data collected in oscillation mode. In *Methods in enzymology* (eds. C.W.J. Carter and R.M. Sweet). *Academic Press, New York*: 307-326.

Papageorgiou, A.C., Shapiro, R., and Acharya, K.R. 1997. Molecular recognition of human Angiogenin by placental ribonuclease inhibitor - An X-ray crystallographic study at 2.0 Å resolution. *EMBO J.* **16:** 5162-5177.

Pares, X., Nogues, M.V., de Llorens, R., and Cuchillo, C.M. 1991. Structure and function of ribonuclease A binding subsites. *Essays Biochem* **26:** 89-103.

Pellizzari, R.*e.a.* 1996. Structural determinants of the specificity for synaptic vesicle-associated membrane protein synaptobrevin of tetanus and botulinum type B and G neurotoxins. *J. Biol. Chem.* **271:** 20353-20358.

Pizzo, E., Buonanno, P., Di Maro, A., Ponticelli, S., De Falco, S., Quarto, N., Cubellis, M.V., and D'Alessio, G. 2006. Ribonucleases and angiogenins from fish. *J Biol Chem* **281:** 27454-27460.

Pollard, D.R., and Nagyvary, J. 1973. Inhibition of pancreatic ribonuclease A by arabinonucleotides. *Biochemistry* **12:** 1063-1066.

Prabu, M.M., Suguna, K., and Vijayan, M. 1999. Variability in quaternary association of proteins with the same tertiary fold: A case study and rationalization involving legume lectins. *Prot. Struct. Funct. Genet.* **35:** 58-69.

Puhar, A., Johnson, E.A., Rossetto, O., and Montecucco, C. 2004. Comparison of the pH-induced conformational change of different clostridial neurotoxins. *Biochem Biophys Res Commun* **319:** 66-71.

Raines, R.T. 1998. Ribonuclease A. *Chem. Rev.* **98:** 1045-1065.

Ramachandran, G., N., Ramakrishnan, C., and Sasisekharan, V. 1963. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7:** 95-99.

Richards, F.M., Wyckoff, H.W., Carlson, W.D., Allewell, N.M., Lee, B., and Mitsui, Y. 1972. Protein structure, ribonuclease-S and nucleotide interactions. *Cold Spring Harb Symp Quant Biol* **36:** 35-43.

Rickman, C., and Davletov, B. 2005. Arachidonic acid allows SNARE complex formation in the presence of Munc18. *Chem Biol* **12:** 545-553.

Rocha, T.L., Paterson, G., Crimmins, K., Boyd, A., Sawyer, L., and Fothergill-Gilmore, L.A. 1996. Expression and secretion of recombinant ovine beta-lactoglobulin in Saccharomyces cerevisiae and Kluyveromyces lactis. *Biochem J* **313 (Pt 3):** 927-932.

Rollnik, J.D., Karst, M., Fink, M., and Dengler, R. 2001. Botulinum toxin type A and EMG: A key to the understanding of chronic tension-type headaches? *Headache* **41:** 985-989.

Rosenberg, A.H., Lade, B.N., Chui, D.S., Lin, S.W., Dunn, J.J., and Studier, F.W. 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56:** 125-135.

Rosenberg, H.F. 1998. The eosinophil ribonucleases. *Cell Mol Life Sci* **54:** 795-803.

Rosenberg, H.F., Zhang, J., Liao, Y.D., and Dyer, K.D. 2001. Rapid diversification of RNase A superfamily ribonucleases from the bullfrog, Rana catesbeiana. *J Mol Evol* **53:** 31-38.

Rosmann, M.G. 1990. The Molecular Replacement Method. *Acta Crystallogr* **A46:** 73-72.

Rosmann, M.G., and Blow, D.M. 1962. The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallogr.* **15:** 24-31.

Rosseto, O., Schiavo, G., Montecucco, C., Poulain, B., Deloye, G., Lozzi, L., and Shone, C.C. 1994. SNARE motif and neurotoxins. *Nature* **372:** 415-416.

Rummel, A., Karnath, T., Henke, T., Bigalke, H., and Binz, T. 2004. Synaptotagmins I and II act as nerve cell receptors for botulinum neurotoxin G. *J Biol Chem* **279:** 30865-30870.

Russo, A., Acharya, K.R., and Shapiro, R. 2001. Small molecule inhibitors of RNase A and related enzymes. *Methods Enzymol.* **341:** 629-648.

Russo, N., and Shapiro, R. 1999. Potent inhibition of mammalian ribonucleases by 3′,5′-pyrophosphate-linked nucleotides. *J. Biol. Chem.* **274:** 14902-14908.

Russo, N., Shapiro, R., Acharya, K.R., Riordan, J.F., and Vallee, B.L. 1994. Role of glutamine-117 in the ribonucleolytic activity of human angiogenin. *Proc Natl Acad Sci U S A* **91:** 2920-2924.

Russo, N., Shapiro, R., and Vallee, B.L. 1997. 5′-Diphosphoadenosine 3′-phosphate is a potent inhibitor of bovine pancreatic ribonuclease A. *Biochem. Biophys. Res. Commun.* **231:** 671-674.

Samols, D., Thornton, C.G., Murtif, V.L., Kumar, G.K., Haase, F.C., and Wood, H.G. 1988. Evolutionary conservation among biotin enzymes. *J Biol Chem* **263:** 6461-6464.

Satoh, E., Ishii, T., Shimizu, Y., Sawamura, S., and Nishimura, M. 2001. Black tea extract, thearubigiin fraction, counteract the effects of botulinum neurotoxins in mice. *British J. Pharmacol.* **132:** 797-798.

Schiavo, G. 2000. Neurotoxins affecting neuroexocytosis. *Physiol. Rev.* **80:** 717-766.

Schiavo, G., Rosseto, O., and Montecucco, C. 1994. Clostridial neurotoxins as tools to investigate the molecular events of neurotransmitter release. *Semin. Cell Biol* **5:** 221-229.

Schiavo, G., Santucci, A., Dasgupta, B.R., Metha, P.P., Jontes, J., Benfenati, F., Wilson, M.C., and Montecucco, C. 1993. Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct COOH-terminal peptide bonds. *FEBS Lett.* **335:** 99-103.

Schmid, M.F., Robinson, J.P., and DasGupta, B.R. 1993. Direct visualisation of botulinum neurotoxin-induced channels in phospholipid vesicles. *Nature* **364:** 827-830.

Schmidt, T.G., Koepke, J., Frank, R., and Skerra, A. 1996. Molecular interaction between the Strep-tag affinity peptide and its cognate target, streptavidin. *J Mol Biol* **255:** 753-766.

Shaanan, B., Lis, H., and Sharon, N. 1991. Structure of a legume lectin with an ordered N-linked carbohydrate in complex with lactose. *Science* **254:** 862-866.

Shapiro, R., Fox, E.A., and Riordan, J.F. 1989. Role of lysines in human angiogenin: Chemical modification and site-directed mutagenesis. *Biochemistry* **28:** 1726-1732.

Shapiro, R., Riordan, J.F., and Vallee, B.L. 1986. Characteristic ribonucleolytic activity of human angiogenin. *Biochemistry* **25:** 3527-3532.

Shapiro, R., and Vallee, B.L. 1989. Site-directed mutagenesis of histidine-13 and histidine-114 of human angiogenin. Alanine derivatives inhibit angiogenin-induced angiogenesis. *Biochemistry* **28:** 7401-7408.

Shapiro, R., and Vallee, B.L. 1991. Interaction of human placental ribonuclease with placental ribonuclease inhibitor. *Biochemistry* **30:** 2246-2255.

Shone, C.C., Hambleton, P., and Melling, J. 1985. Inactivation of *Clostridium botulinum* type A neurotoxin by trypsin and purification of two tryptic fragments. Proteolytic action near the COOH-terminus of the heavy subunit destroys toxin-binding activity. *Eur J Biochem* **151:** 75-82.

Shone, C.C., Hambleton, P., and Melling, J. 1987. A 50-kDa fragment from the NH_2-terminus of the heavy subunit of *Clostridium botulinum* type A neurotoxin forms channels in lipid vesicles. *Eur J Biochem* **167:** 175-180.

Shone, C.C., Quinn, C.P., Wait, R., Hallis, B., Fooks, S.J., and Hambleton, P. 1993. Proteolytic cleavage of synthetic fragments of vesicle-associated membrane protein, isoform-2 by botulinum type B neurotoxiin. *Eur J Biochem* **217:** 965-971.

Silberstein, S., Mathew, N., Saper, J., and Jenkins, S. 2000. Botulinum toxin type A as a migraine preventive treatment. *Headache* **40:** 445-450.

Simpson, D.M., Alexander, D.N., and O'Brein, C.F. 1996. Botulinum toxin type A in the treatment of upper extremity spasticity: a randomized, double-blind, placebo-controlled trial. *Neurology* **46:** 1306-1310.

Simpson, L., L., and Rapport, M.M. 1971a. The binding of botulinum toxin to membrane lipids: sphingolipids, steroids and fatty acids. *J. Neurochem.* **18:** 1751-1759.

Simpson, L., L., and Rapport, M.M. 1971b. Ganglioside inactivation of botulinum toxin. *J. Neurochem.* **18:** 1341-1343.

Simpson, L.L. 1981. The origin, structure, and pharmacological activity of botulinum toxin. *Pharmacol Rev* **33:** 155-188.

Simpson, L.L. 1986. A preclinical evaluation of aminopyridines as putative therapeutic agents in the treatment of botulism. *Infect Immun* **52:** 858-862.

Skerra, A., and Schmidt, T.G. 2000. Use of the Strep-Tag and streptavidin for detection and purification of recombinant proteins. *Methods Enzymol* **326:** 271-304.

Smith, B.D., and Raines, R.T. 2006. Genetic selection for critical residues in ribonucleases. *J Mol Biol* **362:** 459-478.

Smith, D.B., and Johnson, K.S. 1988. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene* **67:** 31-40.

Smyth, N., Odenthal, U., Merkl, B., and Paulsson, M. 2000. Eukaryotic expression and purification of recombinant extracellular matrix proteins carrying the Strep II tag. *Methods Mol Biol* **139:** 49-57.

Snow, B.J., Tsui, J.K., Bhatt, M.H., Varelas, M., Hashimoto, S.A., and Calne, D.B. 1990. Treatment of spasticity with botulinum toxin: a double-blind study. *Ann Neurol.* **28:** 512-515.

Soncin, F., Strydom, D.J., and Shapiro, R. 1997. Interaction of heparin with human angiogenin. *J Biol Chem* **272:** 9818-9824.

Srinivas, V.R., Reddy, G.B., Ahmad, N., Swaminathan, C.P., Mitra, N., and Surolia, A. 2001. Legume lectin family, the "natural mutants of the quarternary state", provide insights into the relationship between protein stability and oligomerization. *Biochim. Biophys. Acta* **1527:** 102-111.

St Clair, D.K., Rybak, S.M., Riordan, J.F., and Vallee, B.L. 1988. Angiogenin abolishes cell-free protein synthesis by specific ribonucleolytic inactivation of 40S ribosomes. *Biochemistry* **27:** 7263-7268.

Stancombe, P.R., Alexander, F.C., Ling, R., Matheson, M.A., Shone, C.C., and Chaddock, J.A. 2003. Isolation of the gene and large-scale expression and purification of recombinant Erythrina cristagalli lectin. *Protein Expr Purif* **30:** 283-292.

Storoni, L.C., McCoy, A.J., and Read, R.J. 2004. Likelihood-enhanced fast rotation functions. *Acta Crystallogr* **D60:** 432-438.

Stowell, J.K., Widlanski, T.S., Kutateladze, T.G., and Raines, R.T. 1995. Mechanism-based inactivation of ribonuclease A. *J. Org. Chem.* **60:** 6930-6936.

Strydom, D.J., Fett, J.W., Lobb, R.R., Alderman, E.M., Bethune, J.L., Riordan, J.F., and Vallee, B.L. 1985. Amino acid sequence of human tumor derived angiogenin. *Biochemistry* **24:** 5486-5494.

Studier, F.W., and Moffatt, B.A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189:** 113-130.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185:** 60-89.

Sugii, S., and Sakaguchi, G. 1975. Molecular construction of *Clostridium botulinum* type A toxins. *Infect Immun* **12:** 1262-1270.

Sutton, J.M., Wayne, J., Scott, A., O'Brein, S.M., Marks, P.M.H., Alexander, F.C.G., Shone, C.C., and Chaddock, J.A. 2005. Preparation of specifically activatable endopeptidse derivatives of *Clostridium botulinum* toxins type A, B and C and their applications. *Protein Expr Purif* **40:** 31-41.

Suzuki, M., Saxena, S.K., Boix, E., Prill, R.J., Vasandani, V.M., Ladner, J.E., Sung, C., and Youle, R.J. 1999. Engineering receptor-mediated cytotoxicity into human ribonucleases by steric blockade of inhibitor interaction. *Nat Biotechnol* **17:** 265-270.

Svensson, C., Teneberg, S., Nilsson, C.L., Kjellberg, A., Schwarz, F.P., Sharon, N., and Krengel, U. 2002. High-resolution crystal structures of Erythrina cristagalli lectin in complex with lactose and 2'-alpha-L-fucosyllactose and correlation with thermodynamic binding data. *J Mol Biol* **321:** 69-83.

Swaminathan, S., and Eswaramoorthy, S. 2000. Structural analysis of the catalytic and binding sites of Clostridium botulinum neurotoxin B. *Nat Struct Biol* **7:** 693-699.

Taylor, C.B., Bariola, P.A., delCardayre, S.B., Raines, R.T., and Green, P.J. 1993. RNS2: a senescence-associated RNase of Arabidopsis that

diverged from the S-RNases before speciation. *Proc Natl Acad Sci U S A* **90:** 5118-5122.

Teneberg, S., Angstrom, J., Jovall, P.A., and Karlsson, K.A. 1994. Characterization of binding of Gal beta 4GalNAc-specific lectins from *Erythrina cristagalli* and *Erythrina corallodendron* to glycosphingolipids. Detection, isolation, and characterization of a novel glycosphinglipid of bovine buttermilk. *J. Biol. Chem.* **269:** 8554-8563.

Tokunaga, M., Ishibashi, M., Tatsuda, D., and Tokunaga, H. 1997. Secretion of mouse alpha-amylase from Kluyveromyces lactis. *Yeast* **13:** 699-706.

Tonello, F., Schiavo, G., and Montecucco, C. 1997. Metal substitution of tetanus neurotoxin. *Biochem J* **322:** 507-510.

Torfs, H., Shariatmadari, R., Guerrero, F., Parmentier, M., Poels, J., Van Poyer, W., Swinnen, E., De Loof, A., Akerman, K., and Vanden Broeck, J. 2000. Characterization of a receptor for insect tachykinin-like peptide agonists by functional expression in a stable Drosophila Schneider 2 cell line. *J Neurochem* **74:** 2182-2189.

Turton, K., Chaddock, J.A., and Acharya, K.R. 2002. Botulinum and tetanus neurotoxins: structure, function and therapeutic utility. *Trends Biochem Sci* **27:** 552-558.

Turton, K., Natesh, R., Thiyagarajan, N., Chaddock, J.A., and Acharya, K.R. 2004. Crystal structures of Erythrina cristagalli lectin with bound N-linked oligosaccharide and lactose. *Glycobiology* **14:** 923-929.

Vagin, A., and Teplyakov, A. 1997. MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30:** 1022-1025.

Vaidyanathan, V.V., Yoshino, K., Jahnz, M., Dorries, C., Bade, S., Nauenburg, S., Niemann, H., and Binz, T. 1999. Proteolysis of SNAP-25 isoforms by botulinum neurotoxin types A, C, and E: Domains and amino acid residues controlling the formation of enzyme-substrate complexes and cleavage. *J. Neurochem.* **72:** 327-337.

Vallee, B.L., and Auld, D.S. 1990a. Active-site zinc ligands and activated H_2O of zinc enzymes. *Proc Natl Acad Sci U S A* **87:** 220-224.

Vallee, B.L., and Auld, D.S. 1990b. Zinc coordination, function, and structure of zinc enzymes and other proteins. *Biochemistry* **29:** 5647-5659.

Van Damme, E.J.M., Peumans, W.J., Barre, A., and Rouge, P. 1998. Plant lectins: a composite of several distinct familes of structurally and evolutionary related proteins with diverse biological role. *Crit. Rev. Plant Sci.* **17:** 575-692.

van den Berg, J.A., van der Laken, K.J., van Ooyen, A.J., Renniers, T.C., Rietveld, K., Schaap, A., Brake, A.J., Bishop, R.J., Schultz, K., Moyer, D., et al. 1990. Kluyveromyces as a host for heterologous gene expression: expression and secretion of prochymosin. *Biotechnology (N Y)* **8:** 135-139.

van Ermengem, E. 1979. Classics in infectious diseases. A new anaerobic bacillus and its relation to botulism [Original publicasion: van Ermengem E (1897) Uebereinen neuen anaeroben bacillus und seine beziehungen zum botulismus. Z Hyg Infektionskr **26**, 1-56.]. *Rev Infect Dis* **1:** 701-719.

Vasconcelos, I.M., and Oliveira, J.T.A. 2004. Antinutritional properties of plant lectins. *Toxicon* **44:** 385-403.

Venkateswarlu, D., and Ferguson, D.M. 1999. Effects of C2′-substitution on arabinonucleic acid structure and conformation. *J. Am. Chem. Soc.* **121:** 5609-5610.

Vescia, S., Tramontano, D., Augusti-Tocco, G., and D'Alessio, G. 1980. In vitro studies on selective inhibition of tumor cell growth by seminal ribonuclease. *Cancer Res* **40:** 3740-3744.

Vriend, G. 1990. A molecular modeling and drug design program. *J. Mol. Graph.* **8:** 52-56.

Walsh, D.J., and Bergquist, P.L. 1997. Expression and secretion of a thermostable bacterial xylanase in Kluyveromyces lactis. *Appl Environ Microbiol* **63:** 3297-3300.

Walsh, D.J., Gibbs, M.D., and Bergquist, P.L. 1998. Expression and secretion of a xylanase from the extreme thermophile, thermotoga strain FjSS3B.1, in Kluyveromyces lactis. *Extremophiles* **2:** 9-14.

Walz, F.G., Jr. 1971. Kinetic and equilibrium studies on the interaction of ribonuclease A and 2′-deoxyuridine-3′phosphate. *Biochemistry* **10:** 2156-2162.

Wang, W.C., Zinn, K., and Bjorkman, P.J. 1993. Expression and structural studies of fasciclin I, an insect cell adhesion molecule. *J Biol Chem* **268:** 1448-1455.

Weil, L., and Seibles, T.S. 1955. Photooxidation of crystalline ribonuclease in the presence of methylene blue. *Arch Biochem* **54:** 368-377.

Wictome, M., Rossetto, O., Montecucco, C., and Shone, C.C. 1996. Substrate residues N-terminal to the cleavage site of botulinum type B neurotoxin play a role in determining the specificity of its endopeptidase activity. *FEBS Lett* **386:** 133-136.

Wigler, M., Silverstein, S., Lee, L.S., Pellicer, A., Cheng, Y., and Axel, R. 1977. Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. *Cell* **11:** 223-232.

Wilson, K.P., Shewchuk, L.M., Brennan, R.G., Otsuka, A.J., and Matthews, B.W. 1992. Escherichia coli biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains. *Proc Natl Acad Sci U S A* **89:** 9257-9261.

Winslow, G.M., Hayashi, S., Krasnow, M., Hogness, D.S., and Scott, M.P. 1989. Transcriptional activation by the Antennapedia and fushi tarazu proteins in cultured Drosophila cells. *Cell* **57:** 1017-1030.

Wu, Y., Mikulski, S.M., Ardelt, W., Rybak, S.M., and Youle, R.J. 1993. A cytotoxic ribonuclease. Study of the mechanism of onconase cytotoxicity. *J Biol Chem* **268:** 10686-10693.

Wyckoff, H.W., Hardman, K.D., Allewell, N.M., Inagami, T., Johnson, L.N., and Richards, F.M. 1967a. The structure of ribonuclease-S at 3.5 Å resolution. *J Biol Chem* **242:** 3984-3988.

Wyckoff, H.W., Hardman, K.D., Allewell, N.M., Inagami, T., Tsernoglou, D., Johnson, L.N., and Richards, F.M. 1967b. The structure of ribonuclease-S at 6 Å resolution. *J Biol Chem* **242:** 3749-3753.

Wyckoff, H.W., Tsernoglou, D., Hanson, A.W., Knox, J.R., Lee, B., and Richards, F.M. 1970. The three-dimensional structure of ribonuclease-

S. Interpretation of an electron density map at a nominal resolution of 2 A. *J Biol Chem* **245:** 305-328.

Xu, M.Q., and Perler, F.B. 1996. The mechanism of protein splicing and its modulation by mutation. *Embo J* **15:** 5146-5153.

Xu, M.Q., Southworth, M.W., Mersha, F.B., Hornstra, L.J., and Perler, F.B. 1993. In vitro protein splicing of purified precursor and the identification of a branched intermediate. *Cell* **75:** 1371-1377.

Xu, Y., and Beckett, D. 1994. Kinetics of biotinyl-5'-adenylate synthesis catalyzed by the Escherichia coli repressor of biotin biosynthesis and the stability of the enzyme-product complex. *Biochemistry* **33:** 7354-7360.

Yamaguchi, H., Yamamoto, C., and Tanaka, N. 1965. Inhibition of protein synthesis by blasticidin S. I. Studies with cell-free systems from bacterial and mammalian cells. *J Biochem (Tokyo)* **57:** 667-677.

Yang, D., Rosenberg, H.F., Chen, Q., Dyer, K.D., Kurosaka, K., and Oppenheim, J.J. 2003. Eosinophil-derived neurotoxin (EDN), an antimicrobial protein with chemotactic activities for dendritic cells. *Blood* **102:** 3396-3403.

Yoon, J.M., Han, S.H., Kown, O.B., Kim, S.H., Park, M.H., and Kim, B.K. 1999. Cloning and cytotoxicity of fusion proteins of EGF and angiogenin. *Life Sci* **64:** 1435-1445.

Young, J.D., Peterson, C.G., Venge, P., and Cohn, Z.A. 1986. Mechanism of membrane damage mediated by human eosinophil cationic protein. *Nature* **321:** 613-616.

Zhang, J., Dyer, K.D., and Rosenberg, H.F. 2000. Evolution of the rodent eosinophil-associated RNase gene family by rapid gene sorting and positive selection. *Proc Natl Acad Sci U S A* **97:** 4701-4706.